A COMPLEX ARCHITECTURE FOR GENETIC MODIFICATION OF CARDIOVASCULAR PHENOTYPES IN MOUSE MODELS OF TGF-BETA VASCULOPATHIES

By

Juan F. Calderón Giadrosic

A dissertation submitted to The Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, MD

March, 2014

Abstract

Our work on mouse models of Loeys-Dietz syndrome (LDS) has shown that these mice recapitulate human disease and develop aortic root aneurysm, which is associated with a signature of increased transforming growth factor β (TGFβ) signaling. As in Marfan syndrome (MFS), aortic disease in LDS mouse models is sensitive to and fully reversed by the action of losartan, an angiotensin receptor blocker (ARB) that blunts TGFβ signaling, highlighting the molecular overlap between these two diseases.

The instrumental role that mouse models of both MFS and LDS have played in furthering our knowledge of the pathophysiology of these diseases illustrates the importance of animal models for the study of human genetic disorders. However, the realization that the genetic background of an animal model can modulate the clinical expression of disease-specific phenotypic features has also highlighted the need for understanding the role of natural genetic variation in phenotypic variability.

The research presented in this dissertation is based on the observation that in the context of LDS, there is great phenotypic variability associated with specific mouse strain backgrounds. We went on to prove that while postnatal aneurysm progression, which is associated with excessive TGFβ signaling, is blunted upon crossing LDS mutations onto a C57BL/6J (B6) background, LDS mice on this background also show high penetrance of perinatal death due to persistent truncus arteriosus with interrupted aortic arch (PTA/IAA), a congenital heart defect previously associated with insufficient TGFβ signaling in the cardiac neural crest.

To assess the effect of the B6 background on the phenotype we performed a test of dominance. A single cross of *Tgfbr2*$^{G357W/+}$ mice congenic on a 129S6/SvEvTac (129SvE) background to B6 resulted in 67% of mutant pups with PTA/IAA, while the control backcross onto 129SvE only resulted in 3.7% (n=1/27) of mutant pups with PTA/IAA ($p<1E-9$, Fisher's exact test). In the context of this mutation, we concluded that the B6 background has a major dominant effect with incomplete penetrance on the generation of PTA/IAA, which could be due to other loci that act in a recessive or dosage-dependent manner. Interestingly, a test of dominance performed with *Tgfbr1*$^{M318R/+}$mice yielded complete absence of outflow tract defects, leading to the conclusion that in this scenario, the B6 background had a recessive effect on the generation of PTA/IAA, which was confirmed when a second backcross onto this background was performed and 50% of the mutant pups showed PTA/IAA.

Since PTA is introduced when LDS mutations are bred onto a mixed background, we used this discrete phenotype to map the relevant modifier alleles. *Tgfbr2*$^{G357W/+}$ mice on a pure 129SvE background were bred to F2 WT mice with extensive recombination between B6 and 129SvE chromosomes. The resulting E17.5 fetuses were phenotyped for TA/IAA and DNA from the phenotyped embryos was collected and genotyped using a SNP array specially designed for mouse intercrosses and linkage studies.

A genome-wide analysis revealed a single major B6-specific locus associated with TA/IAA on mouse chromosome 9, with a $-log_{10}(p)$=9 at a map position coincident with the *Tgfbr2* gene. A minor linkage signal was also apparent on chromosome X with a $-log_{10}(p)$=2.75.

Tgfbr2 emerged as a promising candidate modifier gene on chromosome 9, both by virtue of its known function and the presence of strain-specific variation in the 3'UTR. We recognized that variation in the 3'UTR can affect regulation of gene expression and went on to test for potential strain-specific *Tgfbr2* expression differences. In addition, an unbiased search for microRNAs (miRNAs or miRs) that target *Tgfbr2* revealed that *miR-20b* and *miR-106a* were potential candidates because they are encoded within the suggestive association peak on chromosome X. A luciferase reporter allele harboring the B6 *Tgfbr2* 3'UTR showed dramatically reduced expression when compared to its 129SvE counterpart ($p<1E-9$). Moreover, luciferase activity was equalized between the reporter alleles with strain-specific 3'UTRs upon addition of a miR-106a or *miR-20b* antagonist. Taken together, these data suggest that strain-specific differences in *Tgfbr2* expression are fully accounted for by variable susceptibility to miR-mediated suppression of translation.

To further understand if these changes in gene regulation affect the levels of *TβRII* protein on the surface of cells in the arterial wall and, subsequently, TGFβ activity, we isolated vascular smooth muscle cells (VSMCs) from WT B6 and 129SvE animals and stimulated them with low concentrations of TGFβ1 ligand. Surface levels of TβRII and intracellular levels of phosphorylated SMAD2 (pSMAD2; the major readout of TGFβ pathway activity) were measured by flow cytometry. Higher levels of TβRII and intracellular pSMAD2 were observed VSMCs isolated from 129SvE mice.

While the severity of postnatal phenotypes driven by excessive TGFβ signaling, such as aneurysm progression, is significantly ameliorated on a B6 background when compared to 129SvE in the context of a LDS mutation, the development of prenatal

phenotypes associated with insufficient TGFβ levels is greatly accentuated. We provide evidence for a complex architecture of background-specific phenotypic modification in LDS and perhaps other TGFβ vasculopathies. We show that *Tgfbr2* gene expression is influenced by the levels of microRNAs, such as miR-106a and miR-20b, as well as by the differences in accessibility of these miRs to the 3'UTR of *Tgfbr2* mRNA, which are dictated by sequence variants that distinguish B6 from 129SvE in this gene region.

In summary, we present evidence that suggests a novel regulatory mechanism of TGFβ signaling in the context of LDS-causing mutations. This mechanism is associated with overt defects in heart morphogenesis (increased risk of PTA/IAA) and postnatal aortic wall homeostasis (suppression of postnatal aortic aneurysm progression).

It is our strong conviction that therapeutic strategies that mimic nature's success at attenuating postnatal vascular disease on the B6 background (prominently including TGFβ antagonism) hold strong promise for people with LDS and perhaps other TGFβ vasculopathies.

Advisor: Hal Dietz, M.D.

Reader: Daniel Warren, Ph.D.

Preface

"Caminante, son tus huellas, el camino y nada más,

Caminante, no hay camino, se hace camino al andar".

Antonio Machado

At the moment of writing this dissertation, I think a lot about all the people that made all this possible. In first place, I would like to thank my mentor, Dr Hal Dietz for not only providing all the means required to perform our research at the highest levels but also for teaching me (not painlessly) to think with the utmost scientific rigurosity and to pursue my scientific goals until obtaining irrefutable evidence for, or against the hypothesis of the day. I would also like to thank Hal for his efforts on caring about me (and everyone in the lab) beyond the bench and the walls of our lab.

This, could not have been possible either, had I not counted with the absolute support from many of our Dietz Lab members. In first place, and not by chance, my deepest thanks to Sara Cooke. The very last mission of Sara in the lab has been to be a truly exceptional lab manager which is always pulling tricks to make reagents fly quicker than any courier ever imagined. Before that, she is a patient listener of our whining (including mine, of course) and our own equivalent of the U.N Secretary General, trying to make everyone agree and live peacefully together. I admire her for that. Ultimately, she has been a great friend that is ready to pick the phone at 4 a.m. when we needed to know where the closest E.R was and also to listen to my eternal litanies about American and Chilean politics. Thank you Sara!

Other past and present members of the Dietz Lab have also been determinant in my completion of my Ph.D. David Loch, who tried constantly and fruitlessly to scare me away, turned out to be a great mentor on my first two years in the lab, showing me how to *do things properly*. Thanks also to his wife Christel van Erp, for teaching me how to read his mood and when to run away as quick as we could. Mara Swaim and Yi-Chun Chen for being such great friends in and out of the lab, always ready to help me solve a problem or to get me out of trouble. Liz Gerber has been a terrific baymate, always happy to enlighten us about how fantastic Tom Waits is but also, more importantly, always ready to give us the ultimate technical advice on any experiment that we needed to be successful at once. Elena Gallo (Mc Farlane) has also been a great technical and vocational guide. I believe I speak for the entire lab when I say that we have learned how to love our feisty Italian *dotoressa* but personally, I will be eternally grateful for your advice, your rigurosity and the way you have embraced the role that you play in the lab despite your other much more important responsibilities in life. Thanks a lot Elena!

Other members of the lab that I will deeply miss when I leave are Hamza Aziz (a.k.a The Troublemaker), who has had more than the expected patience towards my pranking and joking and has been a great partner-in-crime the last two years. A tremendous resource in the lab but also a great friend has been Graham Rykiel, who has got deeply involved in this project to an extent that goes far beyond the expected for a summer undergraduate internship. Thanks a lot *Padawan*!

To the whole "new generation" of Dietz Lab's students, namely Suha Bachir, James Beckett, Ben Kang, Nicole Wilson and Shira Ziegler, I wish you the very best in

your careers and I look forward to hearing about your great contributions to science and society.

I owe my deepest acknowledgements and thanks to everyone at our beloved Predoctoral Training Program in Human Genetics and by proxy, at our McKusick-Nathans Institute of Genetic Medicine. To the entire Faculty, because when we were told during our Orientation Day that one of the characteristics in the Program was that everyone was always available to us I did not believe it, and they have proved me wrong over and over again. The learning environment you have created has been the best experience I have ever had. I cannot afford not to mention some Faculty members to whom I am particularly grateful. Dr. David Valle, our Program Director, for believing in me almost six years ago when I insisted over the phone, email and any media available to inquire about the possibility of doing my Ph.D here at Hopkins. Thanks to Dr. Kirby Smith, for always providing the best advice to navigate in the sometimes agitated waters of our Ph.D life (and also for providing some unforgettable memories about Martinis and Mussels!). To Dr. Dan Warren, who has not only become a great scientific counterpart but also a great friend with whom great conversations about soccer, biking and beers can be held at all times.

Special thanks to my thesis committee members, Dr Dimitri Avramopoulos, Dr Aravinda Chakravarti, Dr Andrew McCallion and Dr Roger Reeves. They all played a fundamental role in the progress of my project but also on providing me some of the best technical, professional and personal advice I could have received throughout my life in Graduate School. Finally, to Sandy Muscelli, our Program Administrator, because she developed the patience to understand what I was trying to say and she

played a key role on my involvement in other tasks in the program. Needless to say, she also has been a constant feeder of candy and food leftovers when in need and a mean of transport when snow days cannot be afforded and we needed to be in school on time. You make the program work, Sandy, and I am grateful I was part of this!

To my fellow HumGen students, just a reminder that every time things look dark and obscure, it helps to keep in mind that we are receiving world-class education and, at the same time, we are being paid to do what we love. Doesn't sound like a bad deal, right? Enjoy the journey!

To my family: first, to the one that I can draw a pedigree with. My mom Susana, my dad Ernesto, my step-mom Nedy and all my sibs (Evelyn, Nedy, Tiare and Maximiliano) and all the ones I left behind to come to this strange country in pursue of my dreams, I cannot thank you enough for being so selfless as to let me go away from your day-to-day just to be able to "play with mice" for almost six years. The time we have missed together would not be worth it if I had not counted with your full support from the first time I said "I think I want to go to Hopkins for my Ph.D.". I love you all, and you all own a little bit of this degree.

To my family: the one I have acquired due to "interaction with the environment". Turns out, I came to this country with the idea that I could never have American friends. Today, I am ashamed I thought like that, and I realize that I have acquired a family just like the one I was missing from Chile. Matt (Mateo) Knabel and Molly (and Charley Reece!), Jiji Pandiyan and Seb Morisot, David Gorkin, Nara Sobreira and Nick (Honey!) and Ben (Benny) Leadem and Heather. You have all played such an irreplaceable role on my life here in the U.S. Together, you have all taught me

that a family is not only people that we share a good percentage of alleles with, but also those that are there for you in the good times (yes, that includes officiating weddings!) and the bad ones (yes, this one includes knee surgery(ies)) and I will go back to Chile knowing that we will be family for life.

Last, but definitely not least, to my wife Ana. In these last 10 years we have gone through a lot, and you have always stepped up to follow me and my crazy ideas. I can only hope that you have enjoyed this ride as much as I have, because without your love and support, I could not have done anything of what I have done. You have taught me so much about life, perhaps without even noticing, and you have earned my most profound admiration for what you do, your passion and courage to go out there and be the best wife, daughter, friend and the best biology teacher you can be. I love you and I thank you deeply for all you are in my life.

Baltimore, March 2014

Table of Contents

List of Tables

List of Figures

1. CHAPTER 1. INTRODUCTION

1.1 TGFβ signaling

Transforming growth factor-beta (TGFβ) is a potent cytokine that belongs to a family of dimeric polypeptide growth factors that also includes bone morphogenic proteins (BMPs) and activins. TGFβ can modulate a variety of cellular functions and physiologic processes including cell proliferation, differentiation, apoptosis and synthetic repertoire and regulation of embryonic development, wound healing, and angiogenesis (Blobe et al., 2000). The TGFβ family is highly conserved throughout the entire metazoan subkingdom.

There are three different TGFβ isoforms: TGFβ1, TGFβ2 and TGFβ3. Each is encoded by a different gene that is expressed in a tissue-specific and developmentally-regulated pattern. TGFβ activity can have different consequences depending on context. For example TGFβ can induce differentiation of stem cells but cell cycle arrest in epithelial cells. The number of genes under control of the TGFβ pathway varies from a few in pluripotent stem cells to hundreds in differentiated cells (Massague, 2012).

TGFβ signals trough a tetrameric receptor complex comprised of two type I and two type II receptor subunits. The ligand binds exclusively to the type I receptor subunit (ALK5 or *T*βRI) which is known as the *signal propagating* receptor; the ALK5-ligand complex then binds the type II receptor subunit (TβRII), also known as the *activating* receptor. In the classical or canonical TGFβ-dependent signaling cascade, gene expression is regulated by receptor-mediated activation of SMAD transcription factors, including SMAD2 and SMAD3 (receptor-activated SMADS or R-SMADS), which are

phosphorylated by the Ser/Thr kinase domain in the type II receptors. Phosphorylation of SMAD2 and SMAD3 leads to association with SMAD4 (co-SMAD) and subsequent translocation of this complex into the nucleus, where it regulates expression of many target genes in association with other DNA-binding transcription factors (See Figure 1-1).

Extensive cross-talk exists between this pathway and other signaling pathways. For example, activation of the mitogen-activated kinase (MAPK) pathways by other growth factors can lead to inhibitory phosphorylation of SMAD2 and SMAD3 in the regulatory "linker region", and thus inhibition of signaling propagation (Massague, 2012). Ligand-activated TGFβ receptors can also activate the MAPKs ERK, JNK and p38 (examples of so-called non canonical TGFβ signaling).

1.2 Perturbation of TGFβ signaling in Loeys-Dietz Syndrome and other TGFβ vasculopathies

Loeys-Dietz Syndrome (LDS) is a recently described aortic aneurysm syndrome with autosomal dominant inheritance and systemic manifestations that include aggressive aneurysms throughout the arterial tree, arterial tortuosity, hypertelorism, bifid uvula and cleft palate. Other common manifestations include clubfoot deformity, craniosynostosis and congenital heart defects in the form of bicuspid aortic valve (BAV), atrial septal defect (ASD) and patent ductus arteriosus (PDA) (Loeys et al., 2005). LDS shows significant phenotypic overlap with another connective tissue disorder called Marfan syndrome (MFS), which is caused by heterozygous mutations in the gene encoding the extracellular matrix protein fibrillin-1 (*FBN1*) (Dietz et al., 1991).

LDS is caused in the majority of cases by heterozygous mutations in the genes encoding either the type I (*TGFBR1*) or type II (*TGFBR2*) TGFβ receptor subunits. Less commonly, LDS is caused by heterozygous mutations in the genes encoding SMAD3 (an intracellular mediator of signaling) or the TGFβ2 ligand. LDS belongs to a class of connective tissue disorders (CTDs) that are caused by mutations in genes that code for different effectors or regulators of the TGFβ signaling pathway, unified by a tissue signature for high TGFβ signaling, referred to as the TGFβ vasculopathies. This list includes Marfan Syndrome (MFS), Shprintzen-Goldberg Syndrome (SGS), Arterial Tortuosity Syndrome (ATS), Bicuspid Aortic Valve with Aneurysm (BAAV), some forms of Ehlers-Danlos Syndrome (EDS) and Recessive Cutis Laxa. See Table 1 for a list of the genes altered in these diseases.

The role of TGFβ in the development and progression of aortic aneurysm has been under intense scrutiny. Studies done with tissue samples from patients with MFS, LDS and other disorders that present with aortic aneurysm show a consistent signature for high TGFβ signaling, as evidenced by increased phosphorylation of SMAD proteins (SMAD2 and SMAD3) as well as increased phosphorylation and activation of signaling molecules belonging to the so called non-canonical TGFβ pathway (ERK1/2, JNK and p38)(Habashi et al., 2011; Holm et al., 2011; Neptune et al., 2003). Intriguingly, LDS-causing mutations almost always substitute conserved amino acids in the serine/threonine kinase domain of either TGFβ receptor subunit, and render the mutant receptors unable to propagate signaling when expressed in cells naïve for the corresponding subunit (Mizuguchi et al., 2004). More recently, heterozygous loss-of-function mutations in *SMAD3* or *TGFB2* have also been associated with syndromic presentations of thoracic

aortic aneurysm and have been designated as new subtypes of LDS, LDS3 and LDS4, respectively (Lindsay et al., 2012; van de Laar et al., 2011). The fact that loss-of-function mutations in positive regulators of TGFβ signaling lead to syndromes associated with a tissue signature of excessive and not, as one might expect, defective signaling has led to great controversy in the field regarding the precise role of TGFβ signaling in aneurysm development and the wisdom of therapeutic strategies aimed at TGFβ antagonism. Analysis of tissue samples obtained from patients or mouse models has clearly shown that high TGFβ signaling correlates with the development and progression of aortic aneurysms and many of the features of MFS, LDS and other TGFβ vasculopathies. However, other phenotypic features of LDS, most prominently cleft palate, have a clear association with impaired TGFβ signaling (Bush and Jiang, 2012; Proetzel et al., 1995). The cause of these paradoxical findings has not been fully elucidated, but it is likely that it will rest in the complexity of the interactions between different cell types in each particular context.

1.3 Perturbation of TGFβ signaling and Congenital Heart Defects

Embryonic development of the heart is a process in which cell lineages of diverse origins contribute to the different structures of this organ. While most of these structures are derived from the mesoderm, others, like the cushions that will form the outflow tract (OFT) and the pharyngeal arches, are also composed by ectoderm-derived cardiac neural crest cells (CNCCs) and by cells from the second heart field (SHF) (Mjaatvedt et al., 2001). Formation of the heart tube and its subsequent looping to form a four-chambered structure involves the fine regulation of several signaling cascades that will regulate

proliferation, migration and differentiation of many cell types in highly specific patterns. CNCCs populate the pharyngeal arch arteries early in heart development (between day E8 and E10 of mouse embryonic development) where they will form the aorticopulmonary septum, which is the structure that will divide the OFT into the aorta and the pulmonary artery (Jiang et al., 2000). This process involves a very well-orchestrated process of proliferation and migration of the different cell types involved, and is therefore highly sensitive to changes in the molecules that regulate these processes, prominently including those in the TGFβ and the BMP pathways (Marvin et al., 2001; Schneider and Mercola, 2001).

In mice, homozygous deletion of *Tgfb1* or *Tgfb2* causes perinatal lethality due to a vast array of heart malformations. Cell specific manipulations of TGFβ signaling in CNCCs by homozygous deletion of either *Tgfbr1* or *Tgfbr2* in this lineage results in deficient septation of the outflow tract, which manifests in the form of patent truncus arteriosus and/or interrupted aortic arch (PTA or IAA, respectively) (Choudhary et al., 2006). Zhou et al showed that SHF-specific deletion of the gene encoding latent TGFβ-binding protein 3 (*Ltbp3*) in zebrafish affects the elongation of the cardiac tube and formation of the precursors of the OFT (Zhou et al., 2011). Deletion of genes encoding other factors in the TGFβ pathway, such as the latent TGFβ-binding protein 1L (*Ltbp1l*), also cause PTA and IAA in mice (Todorovic et al., 2007).

1.4 Modifiers of Disease

Geneticists have long been fascinated by the fact that even "simple" Mendelian traits are influenced by genetic variation at loci distinct from the primary disease locus, a

phenomenon referred to as genetic modification (Bridges, 1919). Often a particular Mendelian disorder is caused by a variety of different mutations in the same gene, called allelic heterogeneity, which can dictate a striking range of phenotypic diversity (Romeo and McKusick, 1994). Adding to this complexity, variation in environmental exposures can also profoundly influence phenotypic expression (environmental modification). Thus, it can be difficult to attribute with certainty the source of clinical modification in human populations that show a diversity of underlying primary disease alleles, variation at other genetic loci and environments. As we will discuss below, the use of animal models of disease can provide powerful tools to dissect this complexity.

1.5 Animal models in the study of genetic diseases

The use of animal models to study the inheritance of phenotypic traits has been documented back to the early 1900's. Gregor Mendel started his studies of inheritance by breeding mice with different coat colors, and only moved on to study pea plants after his Bishop threatened to close the monastery where he resided (Paigen, 2003). By the 1930's an extensive framework had been developed in mice to interrogate important biological processes that ranged from cancer to immunology. Institutions like the Roscoe B. Jackson Memorial Laboratory (currently known as The Jackson Laboratory) were formed specifically to push forward the development of animal models to address all these topics (Paigen, 2003). Mice became the ultimate model organism to study biochemical pathways and mammalian physiology with the advent of genetic maps, first crudely defined by restriction fragment length polymorphisms (RFLPs) but later refined using millions of single nucleotide polymorphisms (SNPs) that distinguish different inbred

mouse strains. The Mouse Genomes Project has made publicly available sequencing data from 17 inbred mouse strains and a comprehensive catalogue of structural and sequence variants across these strains has been built from these data (Yalcin et al., 2012).

The notion that a certain mutation, whether artificially introduced in mice or naturally occurring, could exhibit strong phenotypic variation when bred onto mice from different strains became rapidly evident. One of the earliest examples is the modulation of severity of intestinal obstruction-related death in various strains of mice with a mutation in the *Cftr* gene (mutations in this gene cause cystic fibrosis in humans) (Rozmahe et al., 1996). Another example was the description of the modification of the phenotype of *Apc* mutant mice (colon cancer model) that are intercrossed to mixed backgrounds (Dietrich et al., 1993).

These early findings promoted further analysis of the variation of traits across different strains of mice. Ultimately, this lead to massive efforts to provide researchers with tools that could be used to refine mechanistic understanding of such phenotypes. Among these, the early description of genetic linkage maps of the mouse (Copeland et al., 1993) and the characterization of the genealogy of the different mouse strains (Beck et al., 2000) provided a strong foundation. More recent initiatives include the Collaborative Cross (Churchill et al., 2004) and the development of high-density genotyping arrays specifically designed to detect intercross-specific differences (Churchill, 2007; Yang et al., 2009).

1.6 Figures: Chapter 1

1.6.1 Figure 1-1. The TGFβ pathway

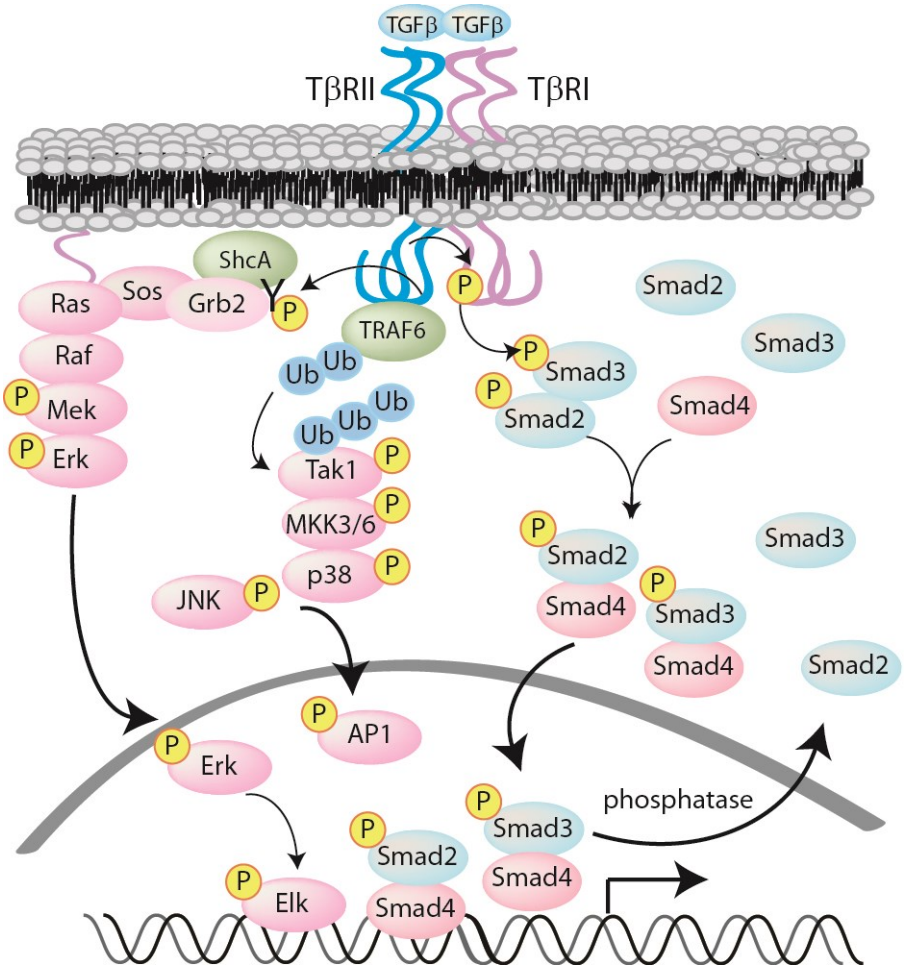


Figure 1-1. TGFβ signaling pathway. The TGFβ receptor is a heterotetramer comprised by TβRI and TβRII (encoded by the genes *Tgfbr1* and *Tgfbr2*, respectively). Activation by TGFβ ligands initiates signaling through the SMAD pathway (canonical TGFβ signaling) or through the MAPK pathways (noncanonical TGFβ signaling).

1.7 Tables: Chapter 1

1.7.1 Table 1-1. Catalogue of Connective tissue disorders referred to as "TGFβ vasculopathies"

Disease	*Gene affected*	*Phenotype MIM #*	*Mode of inheritance*
Loeys-Dietz Syndrome type I	*TGFBR1*	609192	AD
Loeys-Dietz Syndrome type II	*TGFBR2*	610168	AD
Loeys-Dietz Syndrome type III	*SMAD3*	613795	AD
Loeys-Dietz Syndrome type IV	*TGFB2*	614816	AD
Marfan Syndrome	*FBN1*	154700	AD
Shprintzen-Goldberg Syndrome	*SKI*	182212	AD
Ehlers-Danlos Syndrome type IV	*COL3A1*	130050	AD
Arterial Tortuosity Syndrome	*SLC2A10*	208050	AR
Cutis Laxa with Aneurysm	*FBLN4*	614437	AR
Bicuspid AoV with Aneurysm	*NOTCH1* *(others)*	109730	AD

Table 1-1. Abbreviations: AD=Autosomal Dominant

2. CHAPTER 2. CHARACTERIZATION OF THE POSTNATAL EFFECT OF C57BL/6J/129S6/SvEvTac MIXED BACKGROUND ON THE DEVELOPMENT OF AORTIC ANEURYSMS IN LDS MICE

2.1 Introduction

Evidence for the influence of genetic background on the expression of different traits has been documented extensively in several species. In the case of the laboratory mouse (*Mus musculus domesticus*), the effect of different genetic backgrounds on the expression of phenotype in the presence of the same underlying genetic variant has been known for more than 40 years (Coleman and Hummel, 1973; Crawley et al., 1997). Institutions like The Jackson Laboratories and others have devoted great efforts to understanding the genomic architectures of the strains of mice more commonly used in biomedical research and this has resulted in a comprehensive catalogue of the different strains of mice in terms of their potential uses for different studies. Ultimately, initiatives to characterize the genetic variation across the most commonly used strains of mice have generated databases where comparisons between strains at the sequence level can be done using a web-based interface (Keane et al., 2011; Yalcin et al., 2012; Yalcin et al., 2012), facilitating the identification of new genotype-phenotype correlations.

More specifically, modification of cardiovascular phenotypes across different strains of mice has been well documented. In 2009, Wheeler et al. used intercrosses between DBA/2J and C57BL/6J to map a discrete region on mouse chromosome 3 that explained the vast majority of phenotypic variation in a murine model of dilated cardiomyopathy (Wheeler et al., 2009). In regard to modification of TGFβ-related phenotypes, it has been shown that a murine model of MFS displays great variation in the

progression of the aortic root dilatation and subsequent aneurysms when comparing 129SvE and C57BL/6J mice (Lima et al., 2010). We therefore tested the hypothesis that other TGFβ vasculopathies, and more specifically LDS, would display the same phenotypic variation and if so, could be used to further define genetic modifiers of the TGFβ pathway. For this purpose, we used mouse models of LDS that carry point mutations previously observed in LDS patients in either of the TGFβ receptor subunits (*Tgfbr1*$^{M318R/+}$ or *Tgfbr2*$^{G357W/+}$) and that display all the phenotypic features of the disease, including premature death due to aortic dilatation and dissection (Gallo et al., 2014).

2.2 Results

We tested the effect of a single backcross onto C57BL/6J on the progression of aortic disease in LDS mouse models carrying the *Tgfbr1*^{M318R} allele and congenic in a 129S6/SvEvTac background (Gallo et al., 2014). While there was no significant difference between 129S6/SvEvTac congenic and mixed background *Tgfbr1*$^{M318R/+}$mice at 8 or 16 weeks of age, by 24 weeks, 129S6/SvEvTac congenic *Tgfbr1*$^{M318R/+}$ mice showed more severe enlargement of the aortic root as compared to mutant mice in a mixed background ($p<0.007$) (Figure 2-1). An ANOVA test yielded statistical significance for the interaction between backcross status and genotype ($p<1E-8$), and backcross status, genotype status and age ($p<1E-4$).

We then analyzed the aortic root growth rate in these mice. This was calculated by subtracting the aortic root size at 8 weeks of age from the aortic root size at 24 weeks of age. Our results show that LDS mice that are in a mixed background show an aortic root growth rate that is significantly lower than that observed with a congenic 129S6/SvEvTac background ($p<0.0005$) and almost identical to that for wild-type littermates (Figure 2-2).

2.3 Discussion

Enlargement of the aortic root and aortic tear (dissection) are cardinal features of the TGFβ vasculopathies, including MFS and LDS. Murine models for these diseases (previously generated in our lab) have provided significant insights into the pathogenesis of these diseases and have consistently shown a signature of high TGFβ signaling in aortic tissue (Gallo et al., 2014; Lindsay and Dietz, 2011; Neptune et al., 2003). Our results indicate that the C57BL/6J background has a protective effect on the aortic root phenotype of LDS mouse models.

While a protective effect for the C57BL/6J background had been previously proposed in mouse models of MFS through indirect methods to measure the severity of the aortic phenotype (Lima et al., 2010), our results thoroughly document the progression of aortic aneurysms in our murine models of LDS in each strain. Our data shows that different strain backgrounds significantly modify the phenotype of a monogenic disease caused by mutations in a critical component of the TGFβ pathway. Differences in aortic root growth rate become significant by 24 weeks of age and have the strong potential to prevent the catastrophic outcome of LDS (i.e. death due to aortic dissection). .

In subsequent chapters, we will describe the identification of the strain-specific genetic modifiers and the molecular mechanism underlying this modification.

2.4 Materials and methods

2.4.1 Animal models

Animals were housed and experiments were performed with approval by the Johns Hopkins School of Medicine Animal Care and Use Committee. *Tgfbr1*$^{M318R/+}$ and *Tgfbr2*$^{G357W/+}$ LDS mice were previously generated and described (Gallo et al., 2014). Six- to twelve-week old wild-type C57BL/6J (Strain# 000664, The Jackson Laboratory (Jax), Bar Harbor, ME) or 129S6/SvEvTac female mice (Taconic Laboratories) were crossed to *Tgfbr1*$^{M318R/+}$ or *Tgfbr2*$^{G357W/+}$ males congenic on a 129S6/SvEvTac background. The resulting litters were genotyped for the presence of the respective LDS mutations, as described elsewhere (Gallo et al., 2014).

2.4.2 Acquisition of echocardiographic measurements

Echocardiograms were obtained using a parasternal long-axis view and 3 independent measurements of the maximal internal dimension at the sinus of Valsalva on conscious mice whose hair was removed with Nair cream, using the Visualsonics Vevo 660 V1.3.6 imaging system (VisualSonics). Imaging and measurements were performed by a cardiologist who was blinded to genotype and strain background.

2.4.3 Statistical analysis

Statistical analyses were done with the package R (http://www.R-project.org). Data are presented as box-and-whiskers plots. The upper and lower margins of the box represent the 75th and 25th percentiles, respectively; internal line defines the median, and the whiskers define the range. Open circles denote outliers (as defined by the statistical

algorithm), but all values are used in the generation of plots and significance values. In pairwise comparisons, p-values refer to unpaired 2-tailed Student's *t* test. Standard two-way ANOVA analyses to test for different interactions were performed when indicated.

2.5 Figures: Chapter 2

2.5.1 Figure 2-1. C57BL/6J background has a protective effect on the progression of postnatal aortic disease in a murine model of LDS

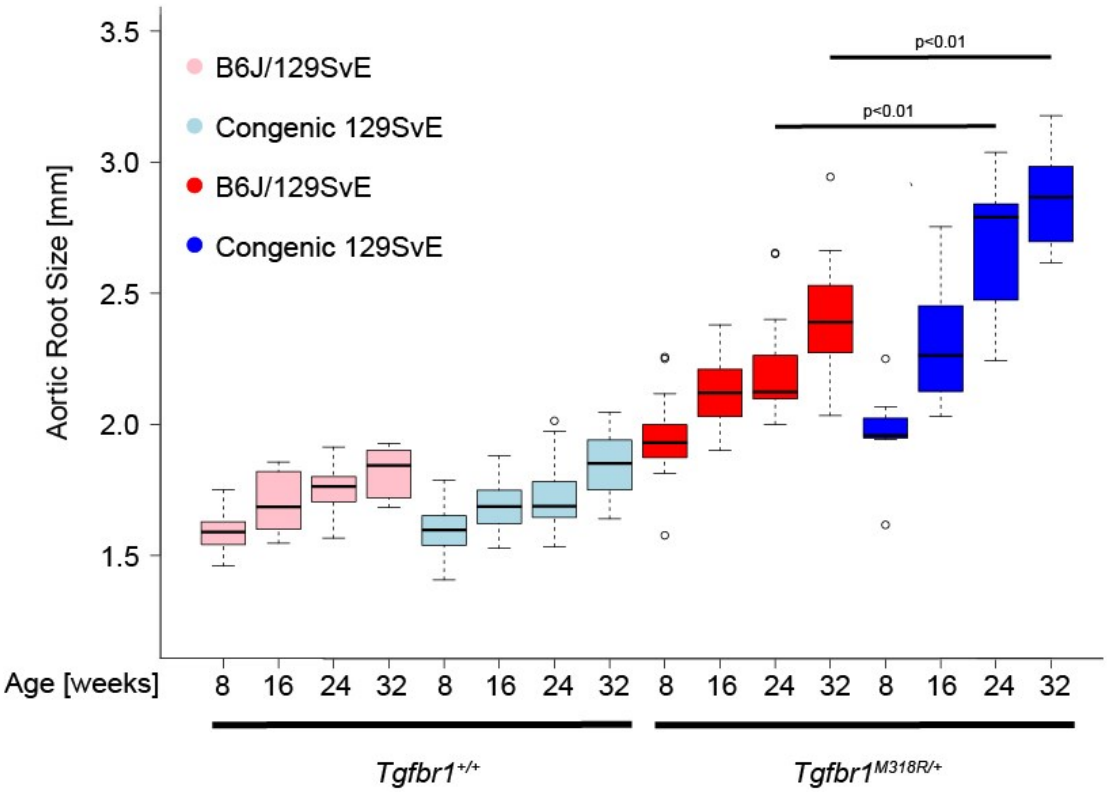


Figure 2-2. Aortic root size was measured at the Sinuses of Valsalva as described in *Methods*. Sample sizes are described in table 2-1 and are at least n=5/group.

2.5.2 Figure 2-2. Aortic root growth Rate of LDS mice congenic on a 129S6/SvEvTac background is significantly higher than LDS mice in a mixed background

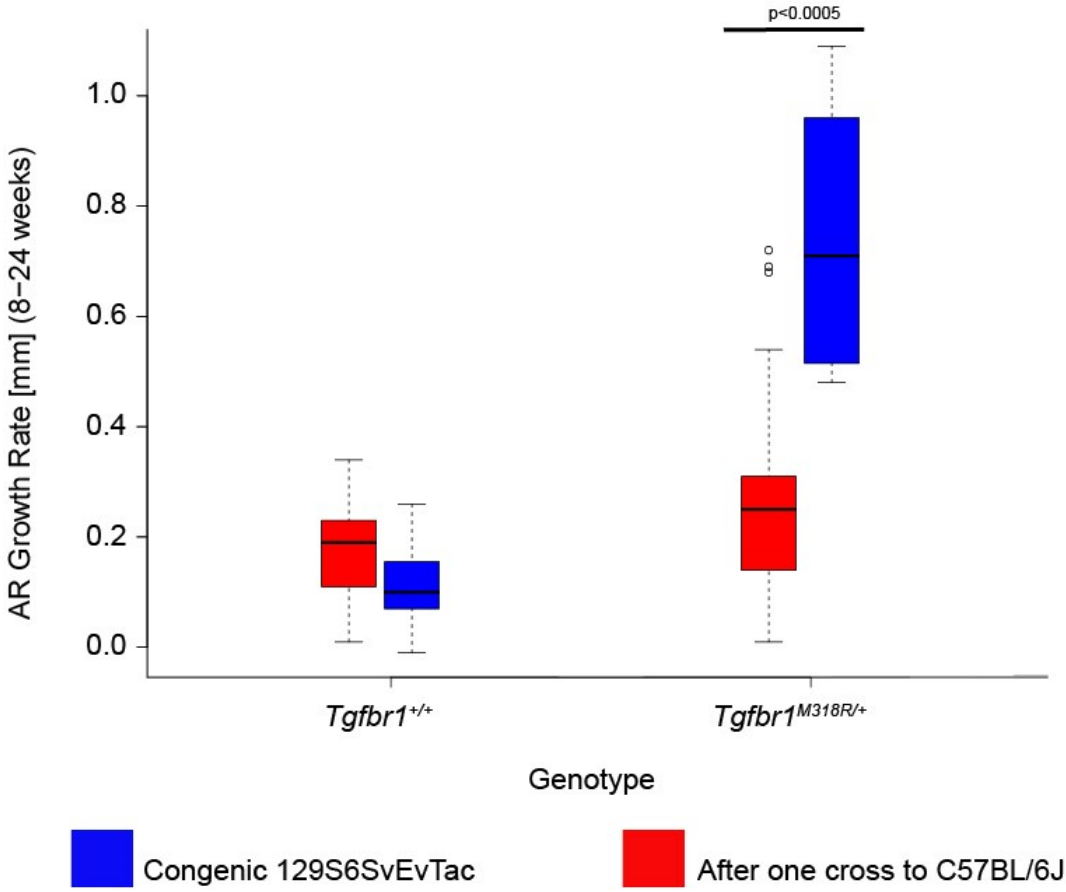


Figure 2-2. Aortic root growth rate was measured as the difference in the aortic root size of an animal at 24 weeks minus the aortic root size at 8 weeks of age (first echocardiographic measurement). Sample sizes are at least n=5 per group.

2.6 Tables

2.6.1 Table 2-1. Sample sizes for echocardiographic measurements of Tgfbr1$^{M318R/+}$ and wild-type littermates in different strains backgrounds

Genotype	Strain	Age (in weeks)	#
Tgbr1$^{+/+}$	Congenic 129SvE	8	27
Tgfbr1$^{M318R/+}$	Congenic 129SvE	8	7
Tgbr1$^{+/+}$	C57Bl/6J/129SvE F1	8	11
Tgfbr1$^{M318R/+}$	C57Bl/6J/129SvE F1	8	20
Tgbr1$^{+/+}$	Congenic 129SvE	16	27
Tgfbr1$^{M318R/+}$	Congenic 129SvE	16	7
Tgbr1$^{+/+}$	C57Bl/6J/129SvE F1	16	12
Tgfbr1$^{M318R/+}$	C57Bl/6J/129SvE F1	16	17
Tgbr1$^{+/+}$	Congenic 129SvE	24	24
Tgfbr1$^{M318R/+}$	Congenic 129SvE	24	6
Tgbr1$^{+/+}$	C57Bl/6J/129SvE F1	24	12
Tgfbr1$^{M318R/+}$	C57Bl/6J/129SvE F1	24	17
Tgbr1$^{+/+}$	Congenic 129SvE	32	18
Tgfbr1$^{M318R/+}$	Congenic 129SvE	32	5
Tgbr1$^{+/+}$	C57Bl/6J/129SvE F1	32	5
Tgfbr1$^{M318R/+}$	C57Bl/6J/129SvE F1	32	9

3. CHAPTER 3. CHARACTERIZATION OF A C57BL/6J-SPECIFIC CONGENITAL HEART DEFECT (CHD) ASSOCIATED WITH LOEYS-DIETZ SYNDROME

3.1 Introduction

Modulation of phenotype due to differences in genetic background across inbred mouse strains has been extensively studied (Churchill, 2007; Nadeau, 2001).

While some authors have directed their efforts to studying the modulation of developmental phenotypes (i.e. CHDs), others have noted that phenotypes expressed in the postnatal stage are also subject to great variation among different strains of mice. Due to the inbred nature of the different strains of mice, these efforts are almost invariably focused on finding regions of the genome that distinguish the strains involved in these studies and trying to correlate this variation to the phenotype of interest. This approach had relative success in the late 1990's with several studies pinpointing "regions of interest" associated with phenotypes that ranged from specific CHDs to more generalized phenotypes such as perinatal lethality.

Among this group, several authors have linked defects in vascular development and perinatal lethality with regions in the mouse genome that modify the function of the TGFβ pathway along with an inherent perturbation in the model (i.e mice null for a component of the TGFβ pathway) leading to the understanding that elements in the mouse genome that vary between different strains negatively affect TGFβ signaling status (Bonyadi et al., 1997; Tang et al., 2003). While the results from these studies indicate that for each phenotype studied there will be one or more regions in the genome that

segregate accordingly, it is not clear whether these regions can be interrogated in the search for a molecular mechanism that directly links the phenotypic variation with any particular effector of the TGFβ pathway.

As mentioned in chapter 2, during our studies on the phenotypic characteristics of our murine models of LDS, we generated mutant mice that were both congenic on a 129S6/SvEvTac background as well as mice that harbored a mixed background between this strain and C57BL/6J. While one backcross onto C57BL/6J had no effect on the expected Mendelian ratio of mutant to wild-type littermates in the progeny of *Tgfbr1*$^{M318R/+}$ mice, one single backcross to C57BL/6J was sufficient to cause 67% perinatal lethality with *Tgfbr2*$^{G357W/+}$ mice due to the presence of septation defects in the OFT such as PTA or IAA. These defects were observed by sacrificing pregnant females at E17.5 days of gestation, in order to allow completion of the OFT septation process (which takes place between E9.5 and 13.5 days of gestation) and avoid perinatal lethality as a confounding factor in our estimations of the penetrance of OFT defects.

The relationship between OFT defects and TGFβ signaling has been highlighted in chapter 1. Here, we show that in the *Tgfbr2*$^{G357W/+}$ LDS mouse model, the presence of OFT defects strictly correlates with the C57BL/6J genomic contribution; we go on to show that penetrance of this trait is high when the *Tgfbr2*^{G357W} mutation is expressed on a mixed background generated by the first backcross of 129SvEvTac onto C57BL/6J (N1), which suggests the presence of at least one modifier locus with a dominant (or pseudodominant) mechanism.

3.2 Results

3.2.1 Patent Truncus Arteriosus (PTA) is a pseudodominant trait in mouse models of Loeys-Dietz Syndrome

The effect of a single backcross (N1) onto C57BL/6J from 129S6/SvEvTac congenic LDS mice, carrying mutations in either subunit of the TGFβ receptor (*Tgfbr1*$^{M318R/+}$ and *Tgfbr2*$^{G357W/+}$) was addressed by performing a test of dominance for the presence of OFT defects (PTA and/or IAA).

0% of E17.5 pups carrying the *Tgfbr1*$^{M318R/+}$genotype presented with OFT defects. However, analysis of *Tgfbr2*$^{G357W/+}$ E17.5 embryos revealed that 67% of them were affected by OFT defects, mainly in the form of PTA (Figure 3-1). This documented dominant inheritance with incomplete penetrance in the context of the *Tgfbr2*$^{G357W/+}$genotype. Results are summarized in table 3-1.

When *Tgfbr1*$^{M318R/+}$male mice resulting from the test of dominance (50% C57BL/6J; 50% 129SvE) were further crossed to wild-type C57BL/6J females, approximately 50% of the mutant E17.5 embryos showed OFT defects (table 3-2), an observation that fits a recessive mode of inheritance of a C57BL/6J-specific modifier allele.

3.3 Discussion

The data generated from our test of dominance supports our prior observation that documented the impossibility of backcrossing any of the LDS-causing mutations previously mentioned onto a C57BL/6J background past the second or third backcross. While the results of the test of dominance performed with *Tgfbr2*$^{G357W/+}$embryos strongly support the hypothesis of a single major dominant effect (as concluded from a penetrance of ≈ 70%), the results observed with *Tgfbr1*$^{M318R/+}$ embryos delineate a recessive nature of modification for the presence of OFT defects. The observation that mutations in the different subunits of the TGFβ receptor interact in differing ways with the C57BL/6J haploid set of chromosomes suggests a complex mechanism of regulation. We hypothesize that PTA requires a reduction of TGFβ signaling below a critical threshold (summarized in figure 3-2). This is achieved with a LDS mutation in one allele of a TGFβ receptor gene (*Tgfbr2*) and a hypomorphic opposing (C57BL/6J) allele of the same gene (figure 3-2A) or a LDS mutation in one allele of a TGFβ receptor gene (*Tgfbr1*) and two hypomorphic (C57BL/6J) alleles of the other TGFβ receptor gene (*Tgfbr2*) (figure 3-2C).

In chapter 6, we will investigate in greater detail the genetic variation between the two strains of interest, with the ultimate goal of revealing potentially novel regulatory mechanisms of the TGFβ signaling pathway which might modulate both pre- and post-natal cardiovascular phenotypes.

3.4 Materials and methods

3.4.1 Generation of E17.5 LDS embryos with mixed genetic background

Animals were housed and experiments were performed with approval by the Johns Hopkins School of Medicine Animal Care and Use Committee. *Tgfbr1*$^{M318R/+}$ and *Tgfbr2*$^{G357W/+}$ LDS mice were previously generated and described by our lab (Gallo et al., 2014). Six- to twelve-week old wild-type C57BL/6J (Strain# 000664, The Jackson Laboratory (Jax), Bar Harbor, ME) or 129S6/SvEvTac female mice (Taconic Laboratories) were timed-mated to *Tgfbr1*$^{M318R/+}$ or *Tgfbr2*$^{G357W/+}$ males congenic on a 129S6/SvEvTac background. Following timed-mating and halothane-induced euthanasia of pregnant females, E17.5 embryos were harvested and their cardiovascular anatomy was observed upon dissection of the abdominal and thoracic walls and injection with approximately 0.5 ml of yellow or blue latex (Ward's Natural Science) into the apex of the left or right ventricle, respectively. Prior to overnight fixation in 10% neutral buffered formalin (Fisher Scientific) and storage in 70% ethanol, tissue samples from every embryo were obtained and stored at -20 Celsius for genomic DNA isolation.

3.4.2 Genotyping of E17.5 embryos

Genomic DNA was obtained from E17.5 embryos by phenol-chloroform purification (Sambrock and Russell, 2001). Determination of LDS genotype and sex was done by PCR. Briefly, amplification of the genomic region harboring the knock-in mutations followed by digestion with *NciI* or *AlwI* demonstrates a diagnostic restriction fragment for the presence of the mutation in *Tgfbr1* or *Tgfbr2*, respectively. Primer sequences can be found in Appendix 1.

Sex of the embryos was determined by multiplex PCR amplification of a genomic region for the *Phex1* and *Sry* genes, which rendered one or two bands for females and males, respectively. Primer sequences can be found in Appendix 1.

3.5 Tables: Chapter 3

3.5.1 Table 3-1. Frequency of OFT Defects in Tgfbr1$^{M318R/+}$ or Tgfbr2$^{G357W/+}$ mice in a mixed background

C57BL/6J wild-type females crossed to:	***Tgfbr1*$^{M318R/+}$**		***Tgfbr2*$^{G357W/+}$**	
E17.5 pup genotype	***Tgfbr1*$^{+/+}$**	***Tgfbr1*$^{M318R/+}$**	***Tgfbr2*$^{+/+}$**	***Tgfbr2*$^{G357W/+}$**
Normal OFT	25 (100%)	43 (100%)	70 (100%)	32 (33%)
OFT defects	0	0	0	65 (67%)

3.5.2 Table 3-2. Frequency of OFT defects in Tgfbr1$^{M318R/+}$ in a second backcross ontoC57BL/6J

C57BL/6J wild-type females crossed to:	**N1 *Tgfbr1*$^{M318R/+}$ males (129S6/SvEvTac/C57Bl/6J)**	
E17.5 pup genotype	***Tgfbr1*$^{+/+}$**	***Tgfbr1*$^{M318R/+}$**
Normal OFT	105 (100%)	55 (49%)
OFT defects	0	58 (51%)

*3.5.3 Table 3-3. Distribution of the incidence of OFT Defects in Tgfbr2$^{G357W/+}$ mice by sex **

C57BL/6J wild-type females crossed to:	***Tgfbr2*$^{G357W/+}$**	
E17.5 pup genotype	***Tgfbr2*$^{G357W/+}$**	***Tgfbr2*$^{G357W/+}$**
Sex	Male	Female
Normal OFT	7	7
OFT defects	16	7

*Table 3-3 only considers sex determination on a subset of embryos due to limited accessibility to genomic DNA from the earliest embryos studied in this experiment.

3.6 Figures: Chapter 3

3.6.1 Figure 3-1. Outflow Tract Anatomy of Tgfbr1$^{M318R/+}$ or Tgfbr2$^{G357W/+}$ mice in a mixed genetic background

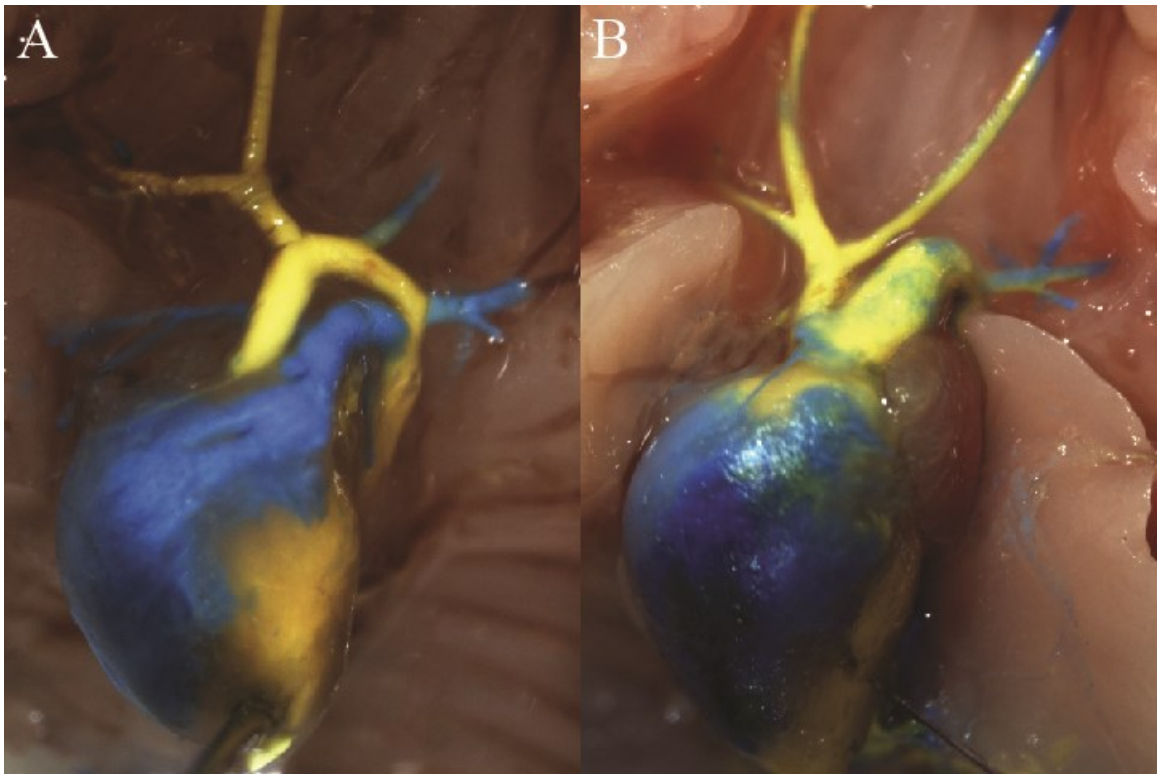


Figure 3-1. A) Bilateral intraventricular injection of latex in heart of a *Tgfbr2*$^{G357W/+}$ E17.5 embryo with normal anatomy shows complete separation of the right and left sides of the circulation with distinct delineation of the aorta and the pulmonary artery arising from the left (yellow) and right (blue) ventricles, respectively. B) *Tgfbr2*$^{G357W/+}$ E17.5 embryo with patent truncus arteriosus (PTA) accompanied by an interrupted aortic arch (IAA). Bilateral intraventricular injection of latex demonstrates the septal defect (intracardiac mixing of colors) associated with PTA/IAA. A patent ductus arteriosus (PDA) is seen in both animals, as expected during fetal life.

3.6.2 *Figure 3-2. TGFβ signaling required in OFT septation is modulated differentially by mutations in different subunits of TGFβ receptor complex and their interaction with C57BL/6J*

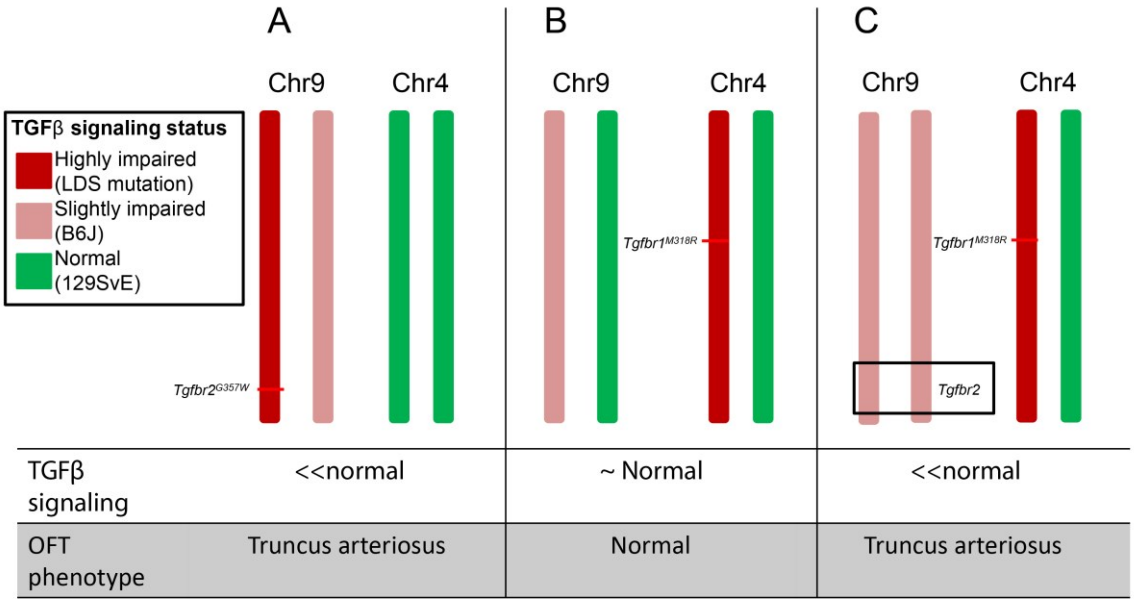


Figure 3-2. Effect of the interaction between LDS-causing mutations in *Tgfbr1* or *Tgfbr2* with C57BL/6J in incidence of OFT defects. A) Interaction of a LDS-causing mutation in *Tgfbr2* with a haploid set of C57BL/6J chromosomes leads to OFT defects (PTA/IAA) (dominance test for $Tgfbr2^{G357W/+}$ mice). B) Interaction of a LDS-causing mutation in *Tgfbr1* with a haploid set of C57BL/6J chromosomes does not lead to OFT defects (PTA/IAA) (dominance test for $Tgfbr1^{M318R/+}$ mice). C) A second backcross of $Tgfbr1^{M318R/+}$ mice onto C57BL/6J is required to observe OFT defects. PTA requires a reduction of TGFβ signaling below a critical threshold that is achieved with a LDS mutation in one allele of a TGFβ receptor gene (*Tgfbr2*) and a hypomorphic opposing (C57BL/6J) allele of the same gene (A) or a LDS mutation in one allele of a TGFβ

receptor gene (*Tgfbr1*) and two hypomorphic (C57BL/6J) alleles of the other TGFβ receptor gene (*Tgfbr2*) (C).

4. CHAPTER 4. MAPPING GENETIC MODIFIERS OF THE CARDIOVASCULAR PHENOTYPE IN LDS MICE

4.1 Introduction

Phenotypic variation among individuals with a similar genetic background (i.e within the same family) has always been of great interest for geneticists. Diseases in which phenotypic features present with incomplete penetrance or variable expressivity certainly are appealing to study modification, assuming interaction between the disease-causing mutation and some other locus or loci in the genome.

Genetic modification is difficult to study using classic human genetic approaches (i.e linkage or association studies looking for alleles that segregate with the specific phenotype of interest). Contributing factors include the outbred nature of most *Homo sapiens* populations, and the confounding influence of allelic heterogeneity and environmental modification, as previously discussed.

As highlighted by Risch and Merikangas (1996), the same limitations apply to the identification of susceptibility loci (or genes) for complex human traits or diseases through classical linkage studies. These authors argue that linkage can only be used when the genes to be identified have a strong effect on the phenotype (i.e. causal gene or major modifiers of the phenotype) and approaches in which many candidate genes can be tested at the same time (i.e. genome-wide association studies) would be limited by the need for large (and often limiting) sample sizes (Risch and Merikangas, 1996).

Given these limitations, analysis of inbred mouse strains offers the opportunity to explore the relationship between the penetrance of a given phenotype and the distribution of alleles of causal and modifying genes (Risch et al., 1993). Typically, in the pre-

genome sequencing era, strains of mice that showed high or low risk for the presence of the phenotype of interest were identified, intercrossed and used to map the susceptibility locus or loci to certain genomic region(s) in linkage disequilibrium with regard to the phenotype of interest. However, given the crude nature of existing linkage maps, it was difficult or impossible to associate loci of modification with mechanisms of modification.

We used a modification of this approach to map a C57BL/6J-specific modifier of the cardiovascular phenotypes in our mouse models of LDS. As discussed in Chapters 2 and 3, we have documented that the C57BL/6J strain interacts with LDS mutations to a) ameliorate postnatal development of aortic root aneurysm and dissection and b) predispose to a developmental cardiovascular phenotype, i.e. PTA.

We chose to use PTA as our trait of interest for the mapping of the critical C57BL/6J modifier(s) because it is: a) a discrete and tractable phenotype that facilitates statistical genetic analyses, b) was shown to be highly specific to the interaction between the C57BL/6J background and LDS mutations, and c) has been previously associated with perturbations of the TGFβ signaling pathway.

4.2 Results

4.2.1 Genome-wide SNP genotyping of E17.5 embryos with PTA and with normal heart anatomy

288 genomic DNA samples were successfully collected and genotyped for a total of 1,449 SNPs corresponding to the entire set of SNPs contained in the Illumina™ Mouse Medium Density Linkage Panel. The distribution of these samples can be seen in table 4-1. A subset of 860 SNPs are polymorphic between C57BL/6J and 129S6/SvEvTac. The mean call rate for this subset was 0.9996 with a Standard Deviation of 0.0014. The genotyping data corresponding to the remaining 589 SNPs were not analyzed due to lack of variation between these two strains (n=540) or technical failure (n=49).

An association analysis with PLINK genetic analysis software (Purcell et al., 2007) included 117 unaffected (normal heart anatomy) and 59 affected (OFT defects) LDS embryos. Generated p-values indicate whether the distribution of genotypes for each SNP is skewed relative to the presence or absence of OFT defects. In Figure 4-2 we represent these in the form of a Manhattan plot, highlighting a strong signal of association located in the distal region of mouse chromosome 9 ($-\log_{10}(p)\approx 9$) and a suggestive signal on mouse chromosome X ($-\log_{10}(p)\approx 2.75$).

A list of the 20 highest ranked SNPs in this analysis is shown in Table 4-2.

4.2.2 Gender-specific effect in the incidence of OFT defects in LDS embryos

Upon the finding of a suggestive association signal in chromosome X we decided to assess whether the incidence of OFT defects correlated with gender. We observed that while 21/174 (12%) of female LDS embryos presented OFT defects, 43/171 (25%) of

males were affected with OFT defects (p=0.002, Fisher's exact test). This difference in the incidence of OFT defects according to sex suggests that the signal observed in the association analysis described in 4.2.1 may be a gender-modified phenomenon.

4.3 Discussion

In this chapter we have described the use of two inbred strains, namely C57BL/6J and 129S6/SvEvTac, and their interaction with LDS mutations, to map susceptibility loci for OFT defects.

As discussed in chapters 3 and 4, this phenotype has been consistently associated with impairment of TGFβ signaling, is highly specific to C57BL/6J and displays a highly dominant effect in its interaction with the *Tgfbr2*^{G357W} allele. In addition, the dichotomous nature of OFTs (PTA or IAA) greatly simplified our association analysis in terms of costs and resources, as compared to the sample size and genotyping resolution that the mapping of a continuous trait (e.g. aneurysm growth rate) would require (Peters et al., 2007).

In our association analysis we observed a signal peak in the distal region of chromosome 9, a region that includes more than 400 genes in a span of approximately 40 Mb. However, *Tgfbr2* was the most prominent candidate gene in the region by virtue of its known function. Based on the genotyping data generated in this cohort of E17.5 LDS embryos, we estimate that in the genomic interval that harbors the SNPs with the highest *–log(p)* scores (chromosome 9), approximately 70% of those embryos with OFT defects carry a C57BL/6J-specific contribution, a result that leads to the conclusion that there may be more than one locus involved.

This notion is reinforced by the finding of a secondary, suggestive signal on mouse chromosome X. Even though this signal does not reach genome-wide significance, it is significantly above the "noise" level and is certainly higher than the uncorrected –log(0.05) value corresponding to the "suggestive" cut off value. As we will further

discuss in chapters 5 and 6, this region also harbors candidate genes that are intimately related to the regulation of TGFβ signaling and that have not previously been studied in the context of cardiovascular phenotypes in mouse models of TGFβ vasculopathies.

4.4 Materials and methods

4.4.1 Animal Models

Animals were housed and experiments were performed with approval by the Johns Hopkins School of Medicine Animal Care and Use Committee. *Tgfbr2*$^{G357W/+}$ LDS mice were previously generated and described in our lab (Gallo et al., 2014). Six- to twelve-week old wild-type C57BL/6J (Strain# 000664, The Jackson Laboratory (Jax), Bar Harbor, ME) or 129S6/SvEvTac female mice (Taconic Laboratories) were crossed to six- to twelve week old wild-type males from a 129S6/SvEvTac background. The resulting litters (C57BL/6J/129S6/SvEvTac F1, or "F1") were bred again in a brother-sister pattern to generate F2 mice with extensive recombination between C57BL/6J and 129S6/SvEvTac chromosomes.

Females from this F2 generation were then mated to *Tgfbr2*$^{G357W/+}$ LDS males congenic on 129SvE in order to generate E17.5 embryos that were phenotyped for the presence of PTA or IAA. Furthermore, some F2 females were mated to pure 129S6/SvEvTac males and the females generated in this cross were bred again to *Tgfbr2*$^{G357W/+}$ 129SvE LDS males to obtain E17.5 embryos with a lower contribution of the C57BL/6J genome (up to 12.5%).

Litters were genotyped for the presence of the respective LDS mutations as described elsewhere (Gallo et al., 2014). For a graphical depiction of the breeding strategy performed refer to figure 4-1.

Following halothane-induced euthanasia of pregnant females, E17.5 embryos were harvested and their cardiovascular anatomy was observed upon dissection of the abdominal and thoracic walls and injection with approximately 0.5 ml of yellow or blue

latex (Ward's Natural Science) into the apex of the left or right ventricle, respectively. Prior to overnight fixation in 10% neutral buffered formalin (Fisher Scientific) and storage in 70% ethanol, tissue samples from every embryo were stored at -20 celsius for further genomic DNA isolation.

4.4.2 Genome-wide SNP genotyping

Genomic DNA obtained from LDS E17.5 embryos (58 with PTA and 117 with normal OFT anatomy) was submitted for genome-wide SNP genotyping at the Center for Inherited Diseases Research (CIDR, Baltimore, MD) via the Genetic Resources Core Facility at The Johns Hopkins University School of Medicine. Briefly, 100 ul of genomic DNA at a concentration of 125 ng/ul was probed on the Illumina™ Mouse Medium Density Linkage Panel, which consists of 1,449 loci (SNPs) optimized for mapping and characterization of intercrosses between different strains of mice.

The Illumina™ GenomeStudio software version 2011.1, Genotyping Module version 1.9.4 (Illumina Inc, San Diego CA) was used as the SNP genotype-calling algorithm. A complete list of the set of markers used in this experiment can be found at http://support.illumina.com/array/array_kits/mouse_md_linkage/downloads.ilmn.

Genotyping data from these samples was used to perform an association analysis as calculated with the PLINK genetic analysis software (Purcell et al., 2007). Briefly, PLINK requires a file with genotyping data for each SNP included in the study as well as phenotypic (with PTA, without PTA) and covariate information (sex), which is comprised in a file with extension *".ped"*; information about the SNPs studied, including chromosome #, rs# and position in bp are incorporated in a file with extension *".map"*.

PLINK was run on these files using the options >*Plink --file [filename] – map3*. Results from this analysis are found in a file with extension *".assoc"* which contains data required to generate a "Manhattan plot" using an open code published by Dr. Stephen Turner (https://github.com/stephenturner/qqman) to be used with the statistical package R (http://www.R-project.org).

4.5 Tables: Chapter 4

4.5.1 Table 4-1. Samples submitted for Genome-wide SNP genotyping

Category	#
Experimental	258
Blind duplicates	12
Internal Controls (CIDR)	12
Investigator Controls	6
Total	**288**

4.5.2 Table 4-2. SNPs with highest scores of association with OFT defects in a mouse model of Loeys-Dietz Syndrome

CHR	*SNP id*	*BP*	*Frequency in Affected*	*Frequency in Unaffected*	*P-value*	*OR*
9	rs3669563	117915299	0.3509	0.08621	9.84E-10	5.73
9	rs6320810	115127089	0.3509	0.09483	5.31E-09	5.16
9	rs13480421	111899990	0.3246	0.1034	3.92E-07	4.165
9	rs3669564	88247268	0.4298	0.2155	3.39E-05	2.744
9	rs3711089	105473772	0.2895	0.1164	6.41E-05	3.093
9	rs13480399	106278521	0.2895	0.1164	6.41E-05	3.093
9	rs3717654	101934506	0.2719	0.1164	0.000272	2.836
9	CEL-9_95875215	95929537	0.2807	0.125	0.000354	2.732
9	rs3689336	96311466	0.2807	0.125	0.000354	2.732
9	rs3700596	86645978	0.2719	0.125	0.000691	2.614
9	gnf09.087.298	90572378	0.2719	0.125	0.000691	2.614
9	rs13480351	94192229	0.2719	0.1293	0.001068	2.515
X	gnfX.044.260	55491587	0.5467	0.337	0.001819	2.372
X	rs13483765	52289226	0.5467	0.3425	0.002435	2.315
X	petM-05810-1	64331236	0.5333	0.3315	0.00259	2.305
X	CEL-X_60181392	63283641	0.5333	0.337	0.003445	2.248
X	rs3157124	66109532	0.5333	0.3425	0.004542	2.194
X	rs13483831	70389009	0.5333	0.3481	0.005938	2.141
X	rs3725966	80159223	0.52	0.3536	0.01347	1.98
X	CEL-X_154048891	157678443	0.5733	0.4088	0.01619	1.943

4.6 Figures: Chapter 4

4.6.1 Figure 4-1. Breeding structure design for mapping modifiers of the cardiovascular phenotype in a mouse model of LDS

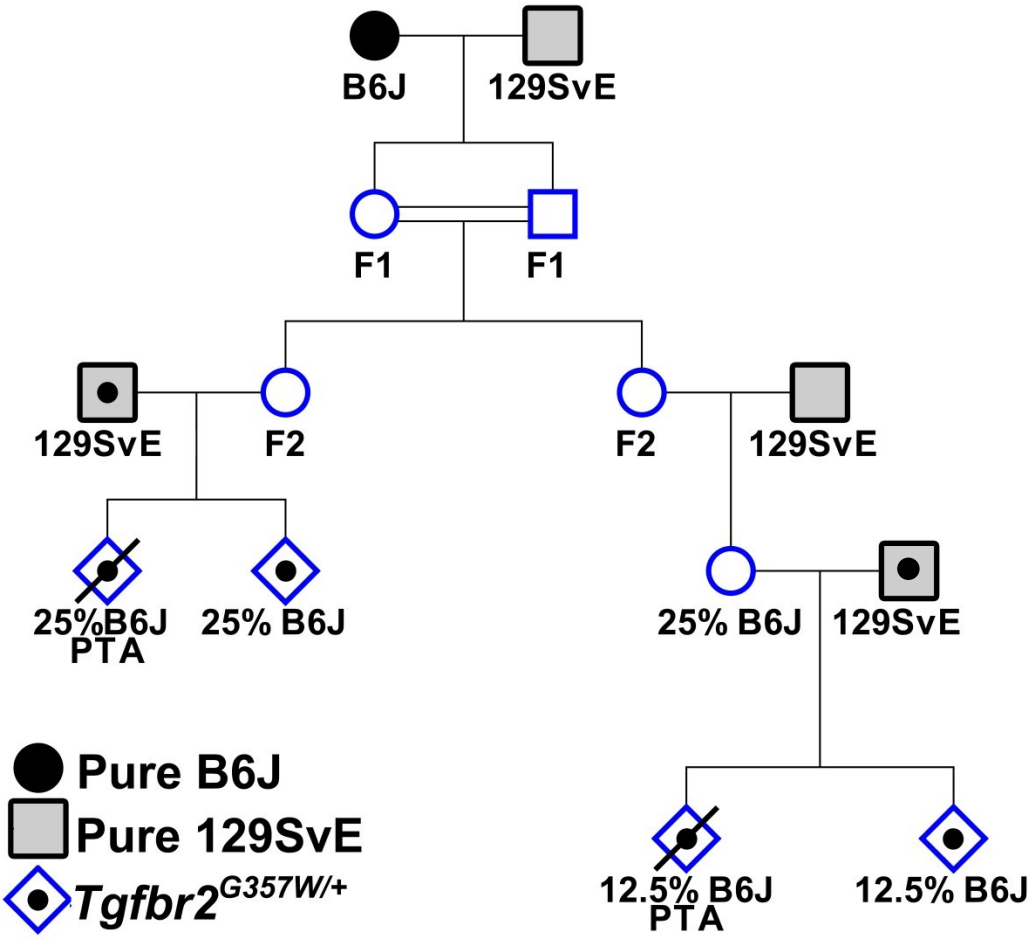


Figure 4-1. Representation of the pedigree structure used to generate a cohort of LDS E17.5 embryos with extensive recombination between C57BL/6J and 129S6/SvEvTac chromosomes. Two groups of *Tgfbr2*$^{G357W/+}$ embryos were used for our analyses, with a degree of C57BL/6J contribution varying from 25% to 12.5%.

4.6.2 *Figure 4-2. Genome-Wide Association analysis of OFT defects in a mouse model of LDS*

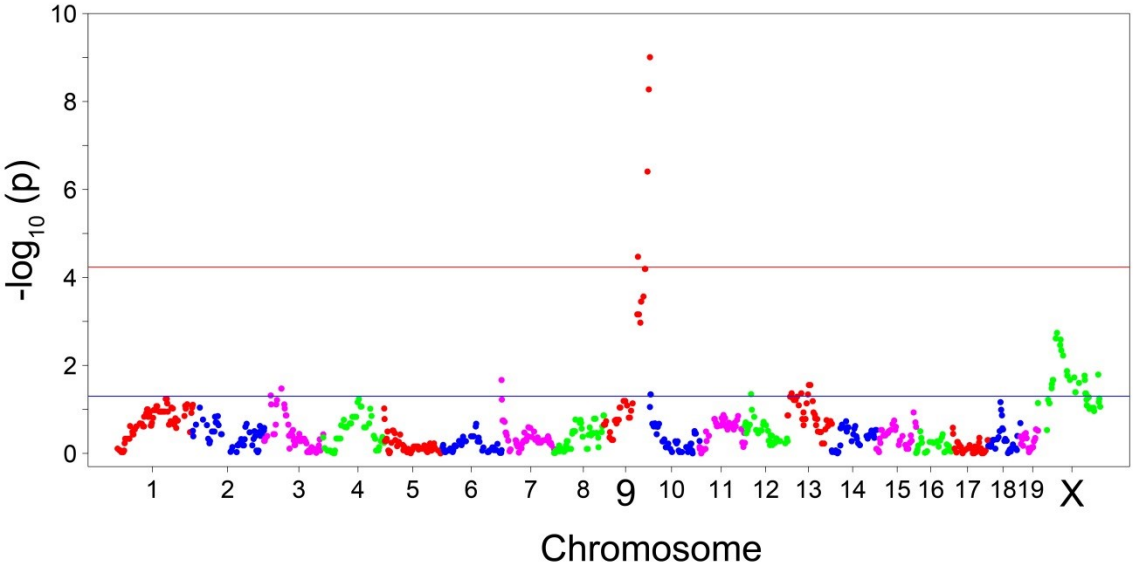


Figure 4-2. Representation of Genome-Wide Association for of OFT defects in *Tgfbr2*$^{G357W/+}$ E17.5 mouse embryos. The $-\log_{10}(p)$ of the calculated p-value for each of the 860 SNPs included in this analysis is plotted as a function of the genomic location of the SNP. Blue line represents $-\log_{10}(0.05)$ or the unadjusted p-value cut-off for significance; red line represents $-\log_{10}(0.05/860)$ or the Bonferroni-adjusted p-value cut-off for genome-wide significance.

5. CHAPTER 5. INTERACTION BETWEEN CANDIDATE GENES IN THE REGIONS OF ASSOCIATION

5.1 Introduction

Classification of traits or diseases according to their mode of inheritance as monogenic, polygenic or complex (the latter implying a strong influence of the environment) was a proper approach during the early efforts for identification of susceptibility locus/loci. When genome-wide screening became available, whether by linkage or association, the observation of multiple signals raised the question of whether these findings implied independent causes of disease (locus heterogeneity) or an interaction between genes at different loci (epistasis) (Vieland and Huang, 2003). While linkage studies generate regions of interest that are never smaller than 2-5 Mb, GWAS hits are located in linkage disequilibrium (LD) blocks that are comparatively shorter ($\leq$ 100 kb) (Ott et al., 2011). However, in both cases, dozens if not hundreds of genes can be found in regions of interest, and without a well-founded list of candidate genes (that is, with strong biological basis for implication in the disease/trait of interest) the likelihood of identifying gene-gene interactions is very low, particularly in the case of quantitative traits.

In the case of qualitative traits, and especially those of dichotomous nature, estimations of penetrance of the trait are used to calculate the effect of each of the loci tested. However, not every statistical model drawn from genotyping data can be matched with a true genetic or biological interpretation, which leaves unanswered the question of whether two loci are truly interacting with one another in the generation of the trait of interest (Sepulveda et al., 2007).

The fact that complex binary traits usually display incomplete penetrance has generally been interpreted as evidence that gene-gene and gene-environment interactions can modify the action of the phenotype-conferring alleles. However, this hypothesis falls short when "pure" genomes and a controlled environment (i.e inbred mice in any modern animal facility) display incomplete penetrance for the development of any given qualitative trait. These observations have prompted the hypothesis of an intrinsic stochastic property in the expression of a genotype at the level of the phenotype (Sepulveda et al., 2007). This can be particularly important in perturbations of morphogenesis, where chance fate decisions in a few progenitor cells can have a profound impact on phenotype.

In this scenario, *epistasis* would have mainly two components: the intrinsic probability of the phenotype-causing allele being expressed and the interaction (if any) of this allele(s) with other alleles or the environment. Several operational definition of epistasis have been proposed: *functional epistasis* relates to the molecular interactions that proteins (and other genetically-determined elements) have with one another; *compositional epistasis*, describes epistasis as the blocking effect of one allele by an allele at another locus (i.e protective or detrimental effect of a second allele); *statistical epistasis* relates to the situation in which "the average deviation of combinations of alleles at different loci is estimated over all other genotypes present within a population" (Phillips, 2008).

In light of these definitions, we aimed at elucidating whether the two loci detected by GWAS in our mouse models of LDS in a mixed background between C57BL/6J and 129S6/SvEvTac displayed epistatic interaction, focusing primarily on *statistical epistasis*

as a way of understanding whether the C57BL/6J contribution to the background of LDS mice with PTA was enriched beyond what could be expected by chance when compared to those LDS mice with a normal heart anatomy.

In this chapter, we show that there is an increased frequency of PTA in those mice harboring C57BL/6J alleles at both loci, that this effect is highly significant, and that while the locus on chromosome 9 is the major modifier, the locus on chromosome X had both a strong cooperative effect and a weak independent effect on the clinical expression of PTA.

5.2 Results

5.2.1 Penetrance of PTA is increased in the presence of C57BL/6J alleles at both chromosome 9 and chromosome X loci in mice with a Tgfbr2^{G357W} allele

If epistatic interaction between the loci on chromosome 9 and X exists, then the penetrance of PTA is expected to be unevenly distributed across the different combinations of C57BL/6J alleles at these loci. That is, the presence of C57BL/6J alleles at both loci should increase penetrance of PTA beyond that expected for each locus by itself. Genotyping data used for GWAS analysis (Chapter 4) was analyzed stratified by phenotype and locus. We find that penetrance of PTA in mice with C57BL/6J alleles at both loci is much higher than expected from our previous observations (test of dominance) and thereafter follows a distribution that documents strong interaction between these two loci.

Indeed, the pairwise comparisons for the penetrance of PTA in the different genotypic classes are significant for almost every pair of categories. Moreover, a general trend test, that seeks to compare the observed results to the null hypothesis that every category has the same probability of showing the phenotype (which is interpreted as lack of interaction between loci) is highly significant ($p<1E-7$), confirming our predictions (Figure 5-1).

We conclude that when the LDS-causing mutation is in the *Tgfbr2* gene and the opposite *Tgfbr2* allele derives from the C57BL/6J background, 47% of mice show PTA. However, this frequency goes up to 87% if this configuration is accompanied by homozygosity (females) or hemizygosity (males) for C57BL/6J alleles at the

chromosome X locus. Indeed, for all genotypic combinations, the penetrance of PTA titrates the combined contribution of the C57BL/6J background at both loci (Figure 5-1).

These observations support the hypothesis that the incidence of PTA is largely dependent on the C57BL/6J contribution on chromosome 9 (*Tgfbr2* locus), but is thereafter strongly modified by the extent of C57BL/6J contribution at chromosome X.

5.2.2 Penetrance of PTA is exacerbated in the presence of C57BL/6J alleles at both chromosome 9 and chromosome X loci in mice with a Tgfbr1^{M318R} allele

On the basis of the results described above for the *Tgfbr2*^{G357W} allele, we reasoned that a similar epistatic interaction between the loci on chromosome 9 and chromosome X should also manifest itself in the case of *Tgfbr1*$^{M318R/+}$ mice after a second backcross onto C57BL/6J which, on average, display a frequency of PTA of approximately 50% (Chapter 3, table 3-2). In fact, we observed that penetrance of PTA is increased among *Tgfbr1*$^{M318R/+}$ mice carrying C57BL/6J alleles at both the *Tgfbr2* locus (chromosome 9) and the locus on chromosome X, with a decreasing penetrance of PTA observed in proportion to the presence of 129S6/SvEvTac alleles at the loci of interest (Figure 5-2).

Following the same principle as in the analysis for mice carrying the *Tgfbr2*^{G357W} allele, mice with the LDS-causing mutation in *Tgfbr1* (which is located in mouse chromosome 4) show PTA at a frequency varying from 61% for those carrying two C57BL/6J alleles at the *Tgfbr2* locus and only C57BL/6J alleles at chromosome X to 12.5% in those mice with one C57BL/6J allele at chromosome 9 and chromosome X.

5.3 Discussion

Our efforts to map regions of the C57BL/6J genome that were associated with PTA in the context of an LDS-causing mutation predicted that one locus on chromosome 9, in the region that contains the *Tgfbr2* gene, was significantly associated with our phenotype of interest. It also confirmed that other regions in the genome (chromosome X) displayed a moderate association that was, however, beyond the noise level and deserved further attention.

As discussed previously, linkage or association studies can yield more than one genomic region as significantly correlated with a phenotype of interest; reasons include locus heterogeneity (in which two or more loci can cause the phenotype independently of each other) or genetic epistasis (in which two or more loci interact in producing the phenotype).

We have demonstrated that in a population with random distribution of C57BL/6J and 129S6/SvEvTac alleles (such as the one used for our GWAS study), there is a significant epistatic interaction between the locus on chromosome 9 and the locus on chromosome X in mice carrying the $Tgfbr2^{G357W}$ allele. The same trend was also observed in mice carrying the $Tgfbr1^{M318R}$ mutation, with the caveat that these mice derive from a second backcross from 129S6/SvEvTac onto C57BL/6J and as such, are 75% C57BL/6J, on average.

The data presented in this chapter reveal that the interaction of these two regions is a *bona fide* example of statistical epistasis. However, it raises important questions in regard to the functional aspects of this relationship and more specifically, how do these

two loci interact to modulate TGFβ signaling to such an extent that the penetrance of PTA shows such extreme variation.

5.4 Materials and methods

5.4.1 Analysis of interaction between loci on chromosome 9 and X in Tgfbr2$^{G357W/+}$ mice

Genotyping data generated for the GWAS previously performed (Chapter 4) was used to designate the genotypes in the SNPs that marked the regions of association for both loci. Genotypes at rs3669563 were used to discriminate between C57BL/6J and 129S6/SvEvTac alleles at the chromosome 9 locus, while genotypes at gnfX.044.260 were used for allelic discrimination at the chromosome X locus.

5.4.2 Analysis of interaction between loci on chromosome 9 and X in Tgfbr1$^{M318R/+}$ mice

SNPs at the chromosome 9 locus (rs13480439) and the chromosome X locus (rs13483770) were genotyped by Taqman™ SNP genotyping (Life Technologies) in a 7900HT platform (Applied Biosystems) using the Allele Discrimination standard protocol. Every genotype call was made with >95% Quality Score over two technical replicates.

5.4.3 Statistical analysis

Statistical analyses were done with the package R (http://www.R-project.org). In pairwise comparisons, p-values refer to Fisher's exact test for a 2x2 contingency table. The trend test was performed using the *prop.trend.test* command which compares the observed frequencies to the expected ones according to scores previously assigned to each genotypic category.

5.5 Figures

5.5.1 Figure -5-1. Segregation analysis of C57BL/6J alleles in E17.5 Tgfbr2$^{G357W/+}$ mice with extensive recombination between C57BL/6J and 129S6/SvEvTac reveals epistasis between chromosome 9 and chromosome X loci

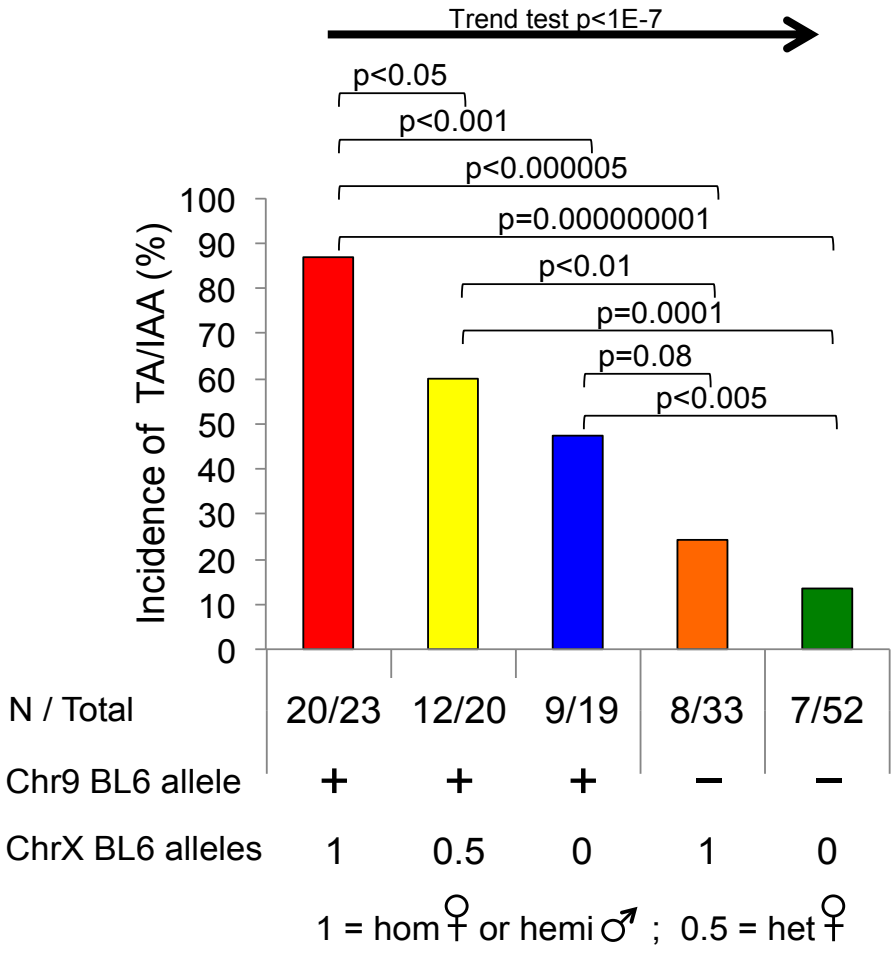


Figure 5-1. Incidence of PTA is primarily determined by the presence of a C57BL/6J allele (opposite to the one carrying the LDS mutation) at the chromosome 9 locus, but is thereafter further modified by C57BL/6J alleles at chromosome X.

5.5.2 *Figure 5-2. Segregation analysis of C57BL/6J alleles in E17.5 Tgfbr1$^{M318R/+}$ mice from a second backcross onto C57BL/6J (from 129S6/SvEvTac) reveals epistasis between loci on chromosome 9 and chromosome X*

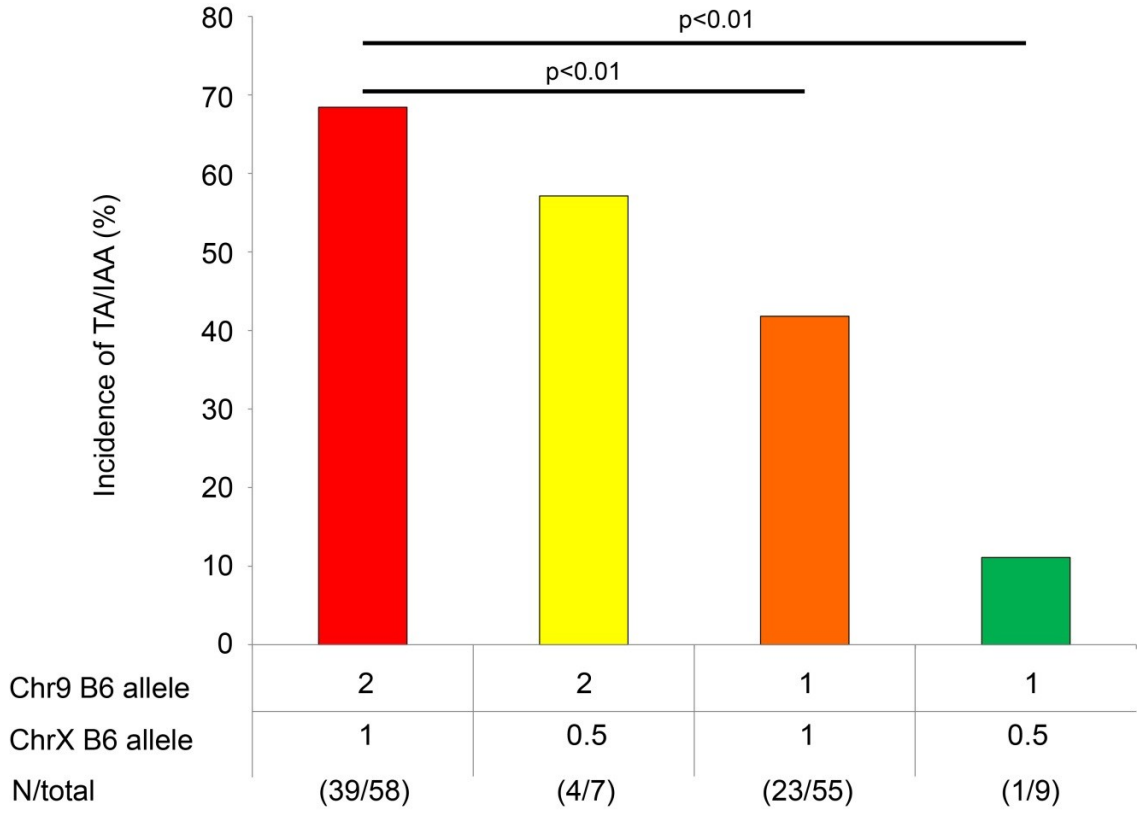


Figure 5-2. Penetrance of PTA in *Tgfbr1*$^{M318R/+}$ mice is also dependent of the interaction between chromosome 9 and chromosome X loci. The presence of C57BL/6J alleles at chromosome 9 is the major driver for occurrence of PTA but is thereafter modulated by the presence of C57BL/6J alleles at chromosome X.

6. CHAPTER 6. STRAIN-SPECIFIC VARIANTS IN THE *TGFBR2* 3'UTR INFLUENCE GENE EXPRESSION

6.1 Introduction

Post-transcriptional regulation of gene expression requires coordination of complex machinery within the cell. It includes addition of the 5' monomethylated CAP structure and subsequent binding of the cap binding complex (CBC) as well as protection of the 3' end of the transcript through binding of the Poly-A binding protein (PABP) complex.

The role of untranslated regions of a transcript (UTRs) has been under intense scrutiny in the last decades, and it is now well understood that variants in these regions can have major effects on gene expression by affecting translation efficiency or mRNA stability, or both. The 5'UTR provides a substrate for the translation initiation complex to bind and scan for the first AUG codon and changes in its sequence can affect this series of events either by disruption of a binding sequence or by formation or modification of secondary structures required for these purposes. An example of how variants in the 5'UTR can affect gene expression and lead to disease states is found in the *BRCA1* gene, which normally works as a tumor suppressor gene. However, a G to C substitution in position -3 in regard to the initiation AUG, which was first described as a somatic mutation found in breast cancer patients, was shown to significantly down-regulate protein translation (Signori et al., 2001).

3'UTR sequences also play a key role in mRNA stability and translation efficiency. It has been reported that mutations that affect termination codons, and therefore the length of the 3'UTR may create unstable transcripts. The polyadenylation

signal (AAUAAA) is another element that, if mutated, can lead to changes in translation efficiency which may result in a phenotype, as in the case of mutations in the polyadenylation signal of the *FOXP3* gene in the context of a syndrome called IPEX (for Immune dysfunction, Poly-endocrinopathy and Enteropathy, X-linked).

Mutations that alter the structure of the 3'UTR can also cause diseases through disruption of interactions with factors that influence mRNA stability, trafficking and translation (Mendell and Dietz, 2001).

Another prominent mechanism of regulation of gene expression related to the 3'UTR involves microRNAs (miRNAs or miRs). These 20-24 bp RNA sequences bind through base-pairing complementarity to sites in the 3'UTR of target genes and repress gene expression through induction of mRNA degradation or (more commonly) inhibition of translation. Mutations in miRNA binding sites have been shown to generate dramatic increases in gene expression with phenotypic consequences, including activation of proto-oncogenes in some cancers (Mayr and Bartel, 2009).

As described in chapter 4, we have identified two major loci associated with PTA in LDS mice, one major locus on chromosome 9 coincident with the genomic location of *Tgfbr2*, and one minor locus on chromosome X. This minor signal centers on the gene encoding the miR-106a~363 locus. Interestingly, this locus contains two microRNAs, *miR-20b* and *miR-106a* that are predicted to target the 3'UTR of *Tgfbr2*.

We hypothesized that molecular interaction between these two miRNAs and the 3'UTR of *Tgfbr2* is at the basis of the epistatic interaction previously demonstrated from a statistical standpoint in chapter 5.

In this chapter, we show that the 3'UTR from the C57BL/6J allele of *Tgfbr2* differs in five positions with respect to the 129S6/SvEvTac allele, and that these differences have functional consequences on the translation efficiency of *Tgfbr2* consistent with our hypothesis that the C57BL/6J background associates with a lower TGFβ signaling environment. Moreover, we show that these differences in translation efficiency are fully attributable to alterations in the activity of *miR-20b* and *miR-106a*.

6.2 Results

6.2.1 Aortic tissue from C57BL/6Jmice express higher levels of primary transcript and mature microRNAs that target Tgfbr2

Given the potential critical role played by the two microRNAs encoded at the chromosome X locus of interest in regulation of the translation efficiency of the *Tgfbr2* transcript, we measured their levels in the two strains of interest by quantitative RT-PCR on RNA extracted from aortic tissue.

Our results indicate that tissue from C57BL/6J has significantly higher levels of the primary transcript encoding both *miR-20b* and *miR-106a.* This trend was also observed when the mature microRNAs were quantitated. In this case, the difference was statistically significant in the case of *mmu-miR-20b* but not in the case of *mmu-miR-106a*, although C57BL/6J also showed slightly higher levels of this transcript (Figure 6-4).

However, we did not observe significant differences in the levels of *Tgfbr2* transcript between the two strains (Figure 6-4), suggesting that levels of receptor expression might only be influenced at the level of translation efficiency and not mRNA stability.

6.2.2 VSMCs derived from C57BL/6J display lower levels of TβRII expression and *TGFβ signaling as compared to 129S6/SvEvTac*

On the basis of our results indicating that expression levels of the miRNAs that inhibit the translational efficiency of *Tgfbr2* are higher in the C57BL/6J, we wanted to test whether this difference has functional consequences on levels of receptor expression and signaling.

Using flow cytometry we were able to simultaneously quantify both receptor expression and levels of SMAD2 phosphorylation (pSMAD2) in primary cultures of vascular smooth muscle cells derived from the ascending aorta of mice of either strain. This analysis was conducted both in the absence and presence of TGFβ1, at different time points.

Our results indicate that VSMCs derived from C57BL/6J aortas display lower levels of TβRII on their surface than their 129S6/SvEvTac counterpart. While there appears to be a slight difference when cells are grown at steady-state in serum, this difference becomes statistically significant at and beyond 15 min of incubation with 1 ng/uL of TGFβ1 ($p<0.01$) (Figure 6-5). TGFβ stimulation leads to acute internalization of receptor complexes at the cell surface, with repopulation of surface expression being largely dependent on *de novo* protein synthesis. The difference in receptor levels associates with lower levels of pSMAD2 in C57BL/6J VSMCs as compared to 129S6/SvEvTac VSMCs, which is most evident after 30 minutes of stimulation with TGFβ1 (Figure 6-5).

6.2.3 *Identification of strain-specific variants in the 3'UTR of Tgfbr2*

Whole-genome next-generation sequencing data have been made publicly available by the Mouse Genome Project at the Wellcome Trust Sanger Institute and a Data Query web-interface (http://www.sanger.ac.uk/sanger/Mouse_SnpViewer/rel-1303) allows users to access Single Nucleotide Polymoprhisms as well as small insertions and deletions and structural variants across 17 strains of mice, including three substrains of 129S and C57BL/6N. A total of 8 variants were found to be polymorphic between

C57BL6N and 129S in the *Tgfbr2* gene according to this source. The distribution of these variants throughout the *Tgfbr2* gene is annotated in Table 6-1 and shown in Figure 6-1. Five variants, which are localized in the 3'UTR region of *Tgfbr2,* were confirmed by Sanger sequencing of genomic DNA from *wild-type* C57BL/6J and 129S6/SvEvTac animals. The two strain-distinguishing SNPs in the coding region of *Tgfbr2* are synonymous, and are therefore unlikely to contribute to functional differences between the strains.

6.2.4 Transcripts carrying Tgfbr2 3'UTR derived from C57BL/6J show reduced translational efficiency as compared to 129S6/SvEvTac

To address the potential effect of the polymorphic variants that distinguish C57BL/6J from 129S6/SvEvTac we introduced the entire 3'UTR from each strain into a reporter vector that encodes firefly luciferase. We then assayed the reporter constructs in NIH/3T3 cells in order to eliminate any potential confounding effect caused by assessing the reporter in cell lines derived from either of the strains being interrogated. We observed that the reporter allele containing the C57BL/6J 3'UTR had approximately 50% less normalized luciferase activity, when compared to the one with the 129S6/SvEvTac 3'UTR ($p<5E-8$) (Figure 6-2).

6.2.5 Tgfbr2 3'UTR from both strains are highly sensitive to alterations on the intracellular levels of mmu-miR-20b and -106a

Upon characterization of strain-specific differences in translation efficiency of a transcript that includes the 3'UTR of *Tgfbr2* from each strain, we decided to investigate

the hypothesis that miRNAs encoded at the *mmu-miR-106~363* locus in mouse chromosome X, which are predicted to target this 3'UTR, can have a significant impact on translational efficiency of our reporter alleles.

We transfected NIH/3T3 cells with mirVana™ inhibitors of mmu-miR-20b or mmu-miR-106a, hypothesizing that inhibiting these miRNAs would increase translation efficiency from either reporter vector. Indeed, we observed increased activity for reporter alleles harboring either the C57BL/6J or 129S6/SvEvTac 3'UTR upon administration of either anti-miR. Of particular note was the full equalization of activity between the reporters with strain-specific 3'UTRs with anti-miR treatment (Figure 6-3).

6.3 Discussion

In Chapter 5 we described a potent interaction between the two loci of interest discovered through our GWAS study and subsequent segregation analysis of C57BL/6J alleles and PTA. Through the series of experiments described in this chapter, we have shown that polymorphic variants in the 3'UTR of *Tgfbr2* that distinguish the two strains of mice have a significant effect on the translational efficiency of this transcript, which can be as much as 50% in our reporter system, with C57BL/6J displaying lower expression.

This difference is abolished when intracellular levels of the microRNAs located at the chromosome X locus are altered, implying that translation levels of *Tgfbr2* are directly modified by *mmu-miR-20b* and *-106a*.

We have also showed that aortic tissue extracted from C57BL/6J mice displays higher expression levels of both primary transcripts and mature *miR-20b* and *miR-106a*. This difference strongly supports our hypothesis that C57BL/6J offers a low TGFβ signaling environment, as previously documented, and that the source of this modification lies in a post-transcriptional modification.

In keeping with our hypothesis, when VSMCs derived from aortic tissue of either strain of mice were stimulated with TGFβ1 ligand we observed that both cell surface levels of TβRII and intracellular levels of pSMAD2 were consistently lower in cells derived from C57BL/6J mice, in agreement with the observation that PTA, a condition triggered by impaired TGFβ signaling, is more penetrant in mice that carry a LDS-causing mutation in the context of increased C57BL/6J contribution at the loci of interest.

We are currently trying to understand the source of this functional variation in the context of the variants localized in the 3'UTR of *Tgfbr2*. Other authors have suggested the relevance of folding structure of mRNAs for the efficiency of its targeting by miRNAs (Long et al., 2007). Our *in silico* mRNA folding studies suggest that the C57BL/6J 3'UTR has a more open conformation in the region of the miR-106a/20b binding site, when compared to the 129SvE 3'UTR, perhaps allowing greater accessibility (Figure 6-6). This would explain how variation in the 3'UTR that is distant from the binding site could influence miR performance.

We have summarized these findings and a model to explain strain-specific phenotypic variation both during embryonic development and postnatal aneurysm progression in LDS in Figure 6-7.

6.4 Materials and methods

6.4.1 Cell culture

3T3 immortalized mouse fibroblasts were grown in DMEM supplemented with 10% fetal bovine serum (FBS), Antibiotic-Antimycotic and L-glutamine, at 37°C and a 5%CO_2 atmosphere.

Primary aortic VSMC cultures were established from the roots and proximal ascending aortas of 10-week-old mice from each strain. Aortas were digested in 300U/ml collagenase II (Worthington Biochemical Corp.) in HBSS (Life Technologies) with Antibiotic-Antimycotic for 20 minutes at 37°and the endothelium and adventitia were removed. Following an overnight incubation in DMEM (Life Technologies) with 20% FBS, aortas were washed with HBSS and digested with 300 U/ml collagenase II and 3U/ml elastase for 45 minutes at 37°C. Aortas were minced into small pieces and incubated with DMEM plus 20% FBS. Seven to ten days were allowed for VSMCs to adhere to the plate, after which the media was changed to DMEM plus 10% FBS.

6.4.2 Molecular Cloning

The 3'UTR region of the *Tgfbr2* gene from C57BL/6J and 129SvE was PCR amplified from mouse genomic DNA (Thermo Scientific, Pittsburgh, PA #F548). Primers used are in Appendix 2. PCR fragments were cloned into an XbaI restriction site in the pGL3-Control Vector (Promega, Madison, WI # E1741) with the In-Fusion®HD Cloning Plus kit (Clontech Laboratories Inc, Mountain View, CA, #638909). Vector sequences were verified by Sanger sequencing (Figure 6-1).

6.4.3 Dual Luciferase Assays

2.5×10^5 NIH/3T3 cells were plated in 24-well plates and transfected 24 hours later with 200 ng of 3'UTR vector (Empty, 3;UTR-C57BL/6J or 3'UTR-129SvE), 5ng of phRL-SV40 (Promega, Madison, WI) and 50 nM of *anti*-mmu-miR-20b-5p or *anti*-mmu-miR-106a-5p or Anti-miR negative control mirVana™ miRNA inhibitor (Ambion® / Life Technologies, Grand Island, NY, #MH10975, #MH12726) using Turbofect (Thermo Scientific, Pittsburgh, PA, #R0531) according to the manufacturer's protocol. For each condition, six biological replicates were performed. 24 hours after transfection, cells were lysed and assayed for *firefly* and *renilla* luciferase activity using the Dual-Luciferase Reporter Assay System (Promega, Madison, WI, #E1960). *Firefly* luciferase activity was normalized to *Renilla* luciferase activity for each transfected well.

6.4.4 Flow Cytometry

VSMCs were grown to approximately 70% confluence in DMEM plus 10% FBS. Cultures were starved in 2% serum for 24 hours before analysis. Cells were lifted from cultures plates by adding ice-cold 0.5 mM EDTA for 10 minutes, washed twice with PBS, and then resuspended in PBS containing 0.5% BSA (Flow Cytometry Staining Buffer). Cells were stained with APC-conjugated anti-mouse TβRII antibody (R&D Systems, FAB532A) or APC-conjugated isotype control antibody (R&D Systems, IC108A) by adding 10 µl of antibody to 10^6 cells in 100 µl. Cells were incubated for 60 minutes on ice in the dark. Unbound antibody was removed by washing the cells in Flow Cytometry Staining Buffer twice. Stained cells were then fixed for 10 minutes at room temperature in 4% PFA in PBS, washed twice with Flow Cytometry Staining Buffer, and

resuspended in 1 ml Flow Cytometry Staining Buffer prior to acquisition. Flow cytometry data were acquired on a BD FACSVerse flow cytometer (BD Biosciences).

Alternatively, an aliquot of these cells was permeabilized overnight in 90% methanol and stained for intracellular levels of phosphorylated SMAD2 by adding 5 µl of antibody (Cat. # 562586, BD Biosciences) to 10^6 cells in 50 µl. Cells were incubated for 60 minutes on ice in the dark. Unbound antibody was removed by washing the cells in Flow Cytometry Staining Buffer twice and resuspended in 1 ml Flow Cytometry Staining Buffer prior to acquisition. Flow cytometry data were acquired on a BD FACSVerse flow cytometer (BD Biosciences).

Results were evaluated, including median fluorescent intensity (MFI) calculations using FlowJo software (Tree Star Inc.)

6.4.5 RNA extraction and quantitative real-time PCR

Aortic tissue was processed using the miRNeasy Mini Kit (Qiagen) plus an automatic bead homogenizer (FastPrep24, MP Biomedicals). For all the extractions, an on-column DNAse digest was performed using an RNAse-free DNAse kit (Qiagen).

cDNA was prepared according to the manufacturer's protocol with a TaqMan High-Capacity cDNA Reverse Transcription kit (Life Technologies). Alternatively, cDNA specific for quantification of mature microRNAs was prepared according to manufacturer's protocol with a TaqMan™ MicroRNA reverse Transcription Kit and primers specific for *mmu-miR-20b* and *-106a*.

Quantitative PCR was performed in three technical replicates per sample with TaqMan™ Universal PCR Master Mix using an Applied Biosystems QuantStudio™ 7 Flex Real-Time PCR System (Life Technologies)

6.4.6 Statistical analysis

Statistical analyses were done with the package R (http://www.R-project.org). Data are presented as box-and-whiskers plots. The upper and lower margins of the box represent the 75th and 25th percentiles, respectively; internal line defines the median, and the whiskers define the range. In pairwise comparisons, p-values refer to unpaired 2-tailed Student's *t* test.

6.4.7 In silico mRNA folding

Predicted folding structures for the stain-specific 3'UTRs were generated using SFOLD software, as previously described (Ding and Lawrence, 2003).

6.5 Tables

6.5.1 Table 6-1. Distribution of variants in the Tgfbr2 gene

Position (GRCm38)	*dbSNP*	*C57BL/6J allele*	*129S6/SvEvTac allele*	*Gene location*
116,175,080	rs33371690	A	G	5' UTR
116,090,516	rs245019337	G	A	CDS, synonymous
116,090,507	rs33371016	G	A	CDS, synonymous
116,090,169	rs30321058	C	T	3' UTR
116,090,000	rs29842530	T	G	3' UTR
116,089,680	rs13468749	C	G	3' UTR
116,089,123	rs33663743	C	A	3' UTR
116,087,843	rs4227974	G	A	3' UTR

6.6 Figures

6.6.1 Figure 6-1. Reporter constructs carrying Tgfbr2 3'UTR from C57BL/6J or 129S6/SvEvTac

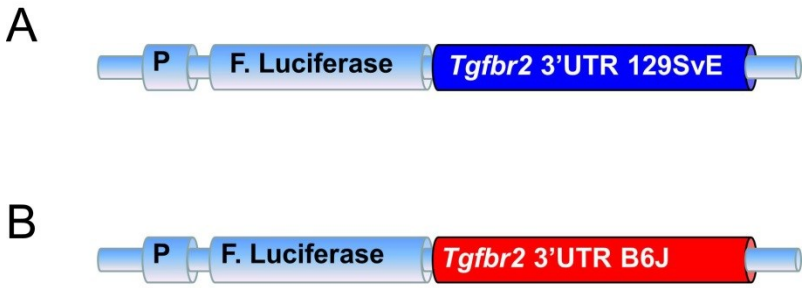


Figure 6-1. Diagram of the reporter vectors used to study regulatory differences between A) C57BL/6J and B) 129S6/SvEvTac.

6.6.2 Figure 6-2 .Assessment of translational efficiency of Tgfbr2 3'UTR from C57BL/6J and 129S6/SvEvTac mice through Dual Luciferase assay

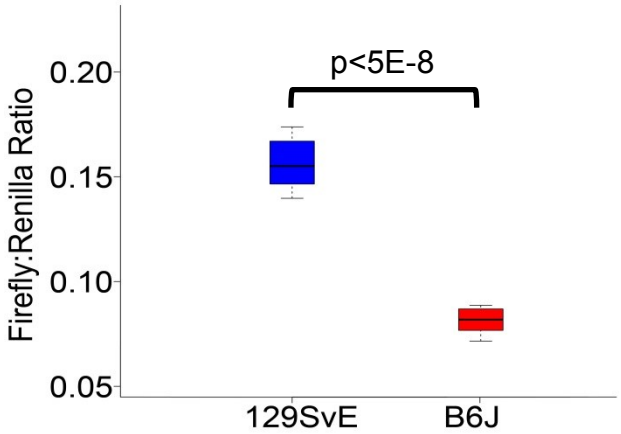


Figure 6-2. NIH/3T3s transfected with a reporter vector carrying the *Tgfbr2* 3'UTR from C57BL/6J display significantly lower luciferase activity than those transfected with a reporter carrying the *Tgfbr2* 3'UTR from 129S6/SvEvTac (p<5E-8).

6.6.3 *Figure 6-3. Sensitivity of Tgfbr2 3'UTR derived from C57BL/6J and 129S6/SvEvTac to alterations in the levels of mmu-miR-20b or mmu-miR-106a*

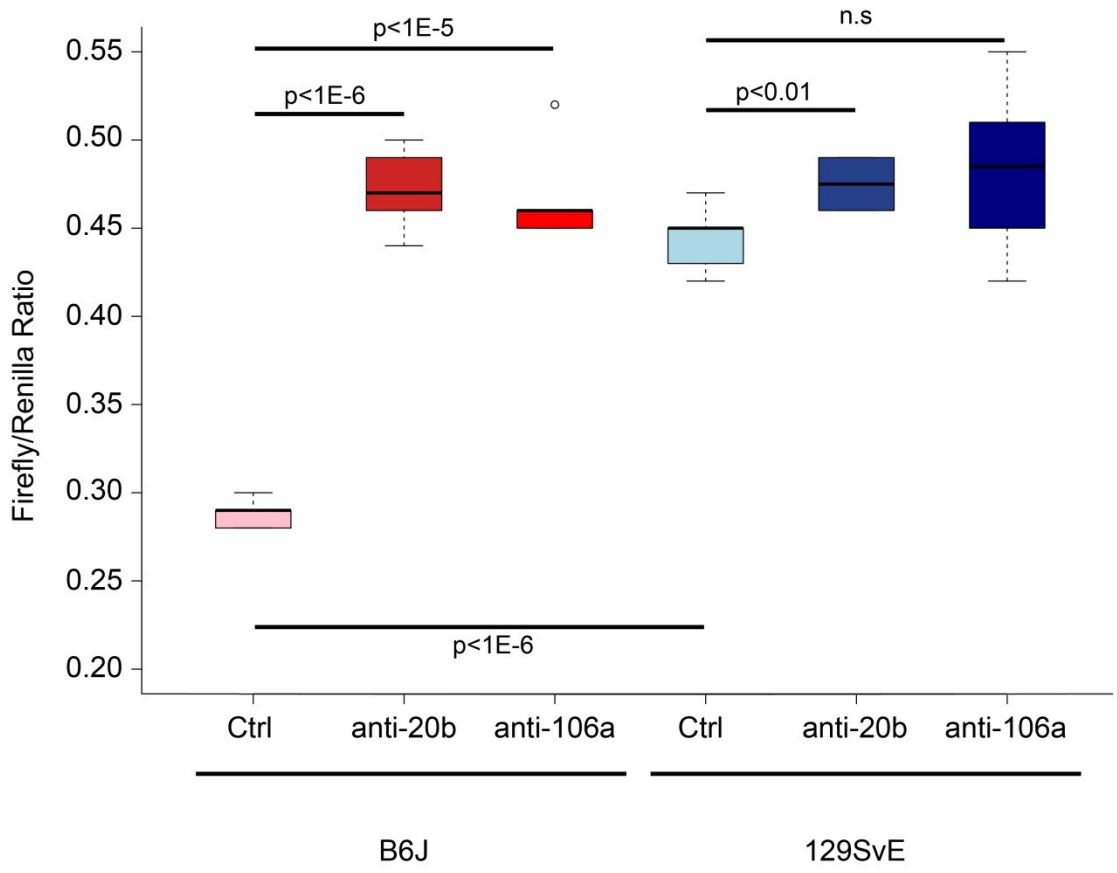


Figure 6-3. NIH/3T3 cells were transfected with reporter vectors carrying the *Tgfbr2* 3'UTR from each of the indicated strains. At baseline, C57BL/6J shows lower levels of reporter translation than 129S6/SvEvTac. Upon addition of inhibitors of either *mmu-miR20b* or *-106a* these levels are equalized between the two strains. P-values were calculated with Student's t-test.

6.6.4 *Figure 6-4. Quantification of RNA levels of candidate genes in the two regions of association with PTA*

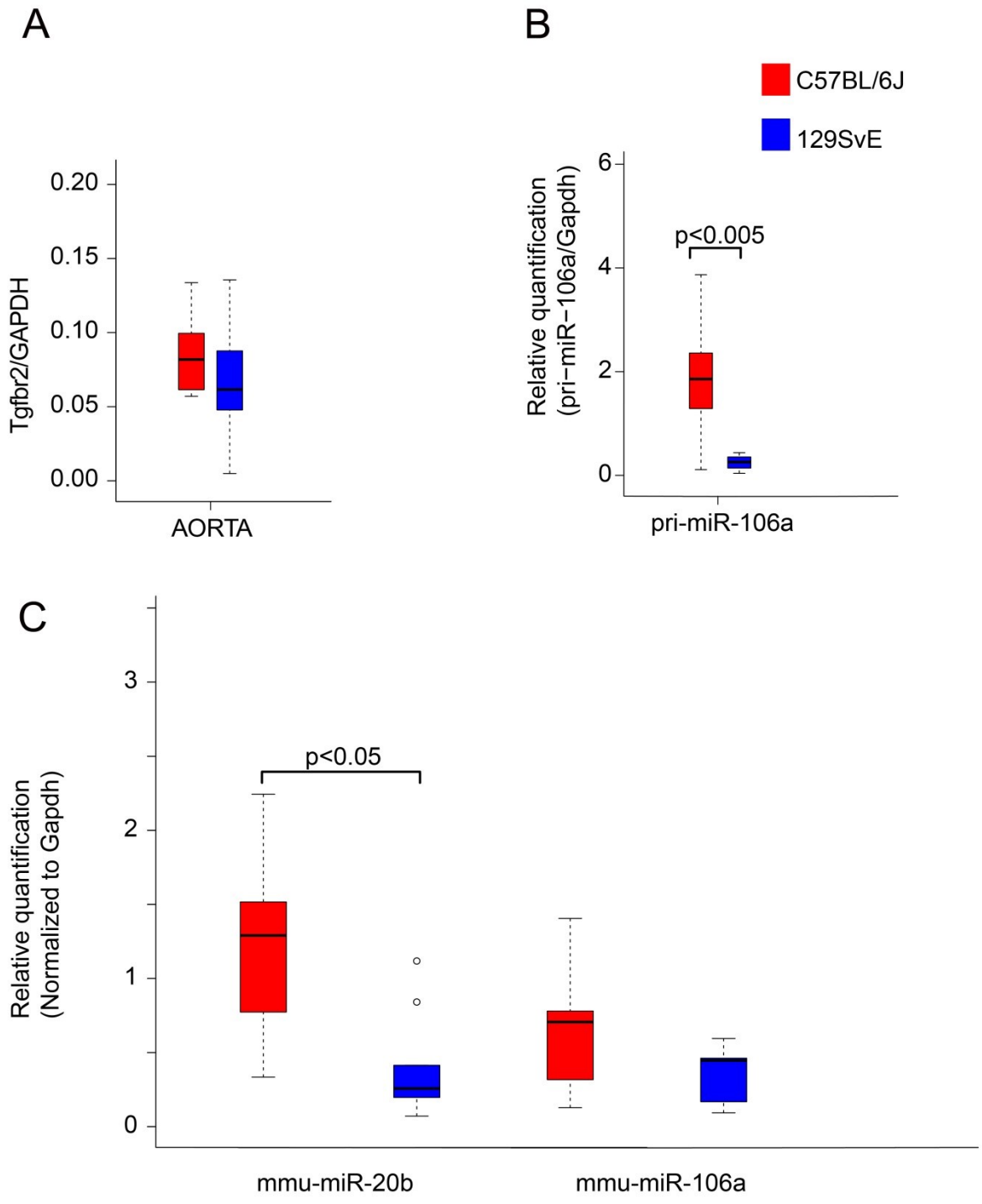


Figure 6-4. RT-PCR quantification of expression levels of candidate genes found in the regions of association in aortic tissue obtained from female mice of either strain (n=3, 3 technical replicates each). A) Quantification of *Tgfbr2* mRNA levels B) Quantification of primary transcript from the *mmu-miR-106~363* cluster. C) Quantification of the mature *mmu-miR-20b* and -*106a* miRNA levels

6.6.5 *Figure 6-5. Levels of surface TβRII and intracellular pSMAD2 in VSMCs derived from C57BL/6J and 129S6/SvEvTac*

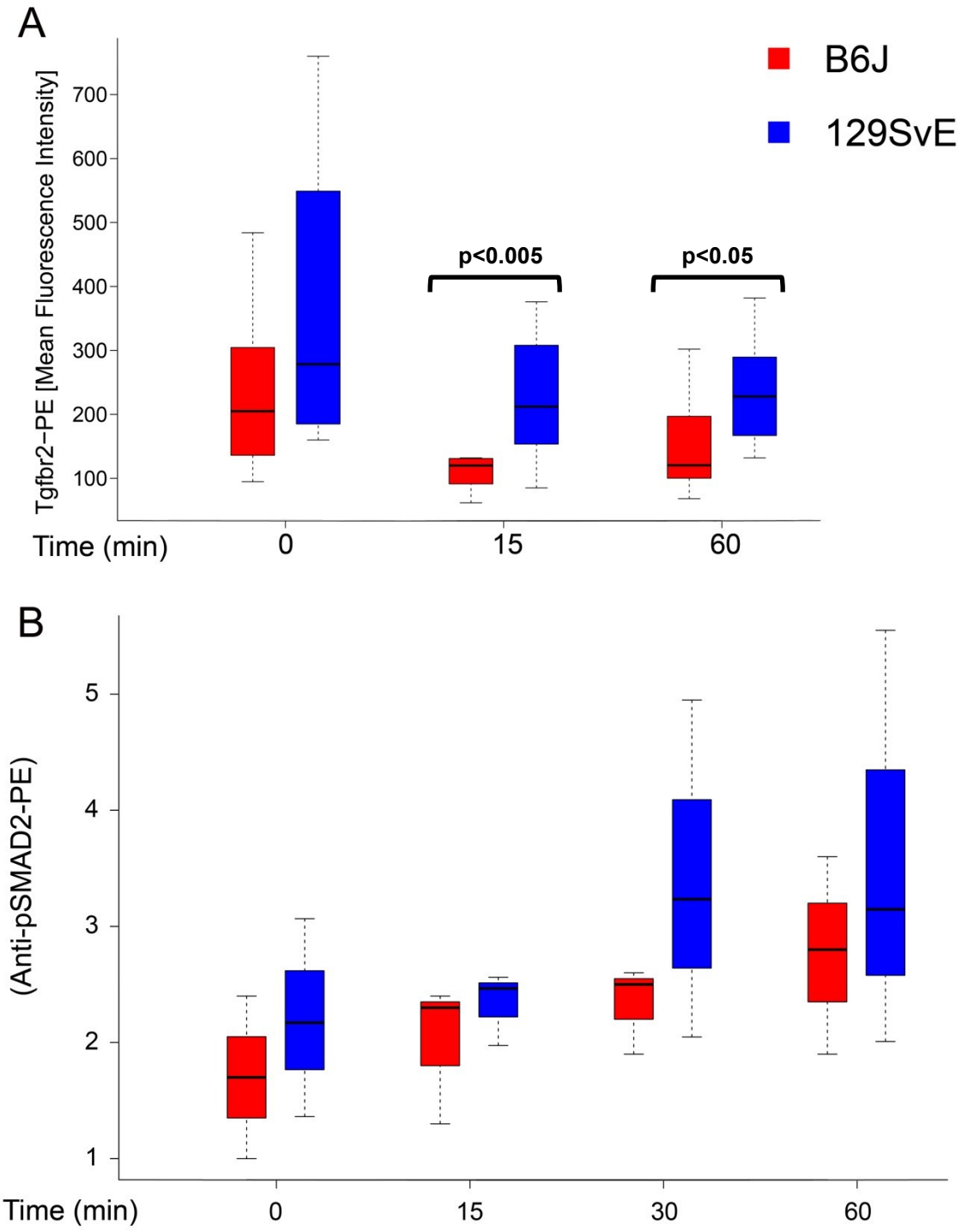
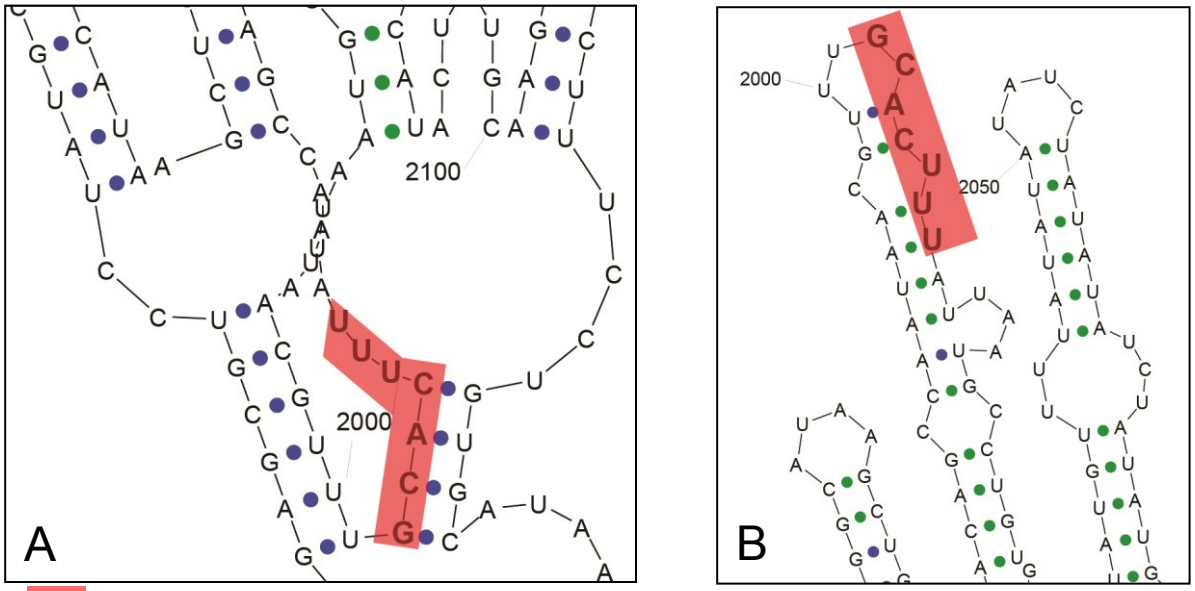


Figure 6-5. Flow cytometric analysis of receptor levels and intracellular signaling of VSMCs derived from each indicated strain in the presence of TGFβ1 for the indicated time points. A) VSMCs from 129S6/SvEvTac display higher levels of TβRII on their surface as compared to C57BL/6J. This difference is significant at 15 and 60 minutes ($p<0.005$ and $p<0.05$, respectively). B) VSMCs from 129S6/SvEvTac display higher intracellular levels of phosphorylated SMAD2 (pSMAD2) as compared to C57BL/6J.

6.6.6 *Figure 6-6. Strain-specific variants in Tgfbr2 influence 3'UTR folding*



Conserved seed sequence for binding of *mmu-miR-20b* and *-106a*

Figure 6-6 Folding structure for *Tgfbr2* full-length mRNA as predicted by SFOLD. A) Location of the seed sequence for *miR-20b* and *-106a* in the 3'UTR of C57BL/6J-specific *Tgfbr2*. B) Location of the seed sequence for *miR-20b* and *-106a* in the 3'UTR of 129S6/SvEvTac-specific *Tgfbr2*.

6.6.7 *Figure 6-7. Modification of cardiovascular phenotypes in LDS mice depends on levels of expression of Tgfbr2 regulatory microRNAs and on strain-specific variation in the Tgfbr2 3'UTR*

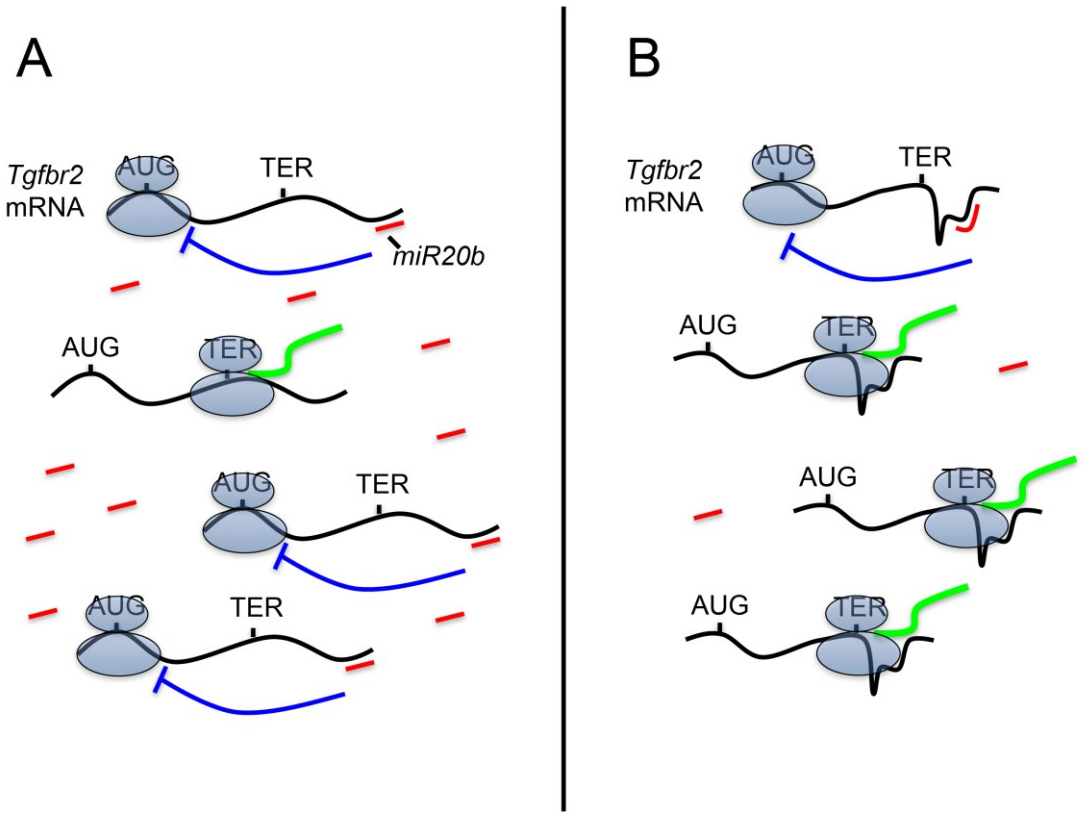


Figure 6-7. *Tgfbr2* expression at the level of translation would be influenced by both expression level and accessibility of microRNAs such as *miR-20b* or *-106a*. A) On the C57BL/6J background, variation on chromosome 9 promotes accessibility to the 3'UTR of *Tgfbr2* and variation on chromosome X dictates high expression of the *miR-106* cluster, lowering expression of the TβRII receptor and TGFβ signaling, inducing OFT defects during embryonic development but ameliorating postnatal progression of aneurysm progression. B) On the 129S6/SvEvTac background, variation on chromosome 9 dictates a less accessible 3'UTR of *Tgfbr2* and lower expression of microRNAs,

culminating in higher expression of the TβRII receptor and TGFβ signaling that promotes exacerbation of postnatal aneurysm progression.

7. CHAPTER 7. CONCLUDING REMARKS

This work demonstrates that the TGFβ signaling pathway can be subject to complex genetic modification with overt phenotypic consequences. Here, we show that strain-specific differences in the TGFβ vasculopathies relates to variation at distinct loci encoded on mouse chromosome 9 (*Tgfbr2* locus) and chromosome X (*miR-106a~363* locus). Moreover, we show robust epistasis between these loci in determining the penetrance of PTA. Given the pseudodominant modification when the LDS mutation is in Tgfbr2, but recessive modification when the mutation is in *Tgfbr1*, our data document a complex architecture for modification that integrates the influence of both the primary disease locus and multiple modifying loci with additive effects ultimately determining the status of TGFβ signaling and hence the predisposition for disease. The different phenotypic outcomes of the interaction between strain-specific alleles of the *Tgfbr2* gene and LDS-causing mutations in regard to OFT septation suggests that TGFβ signaling needs to meet highly specific spatial and temporal thresholds during embryonic development (See Figure 3-2).

Additional complexity is illustrated by our demonstration that variation that dictates low TGFβ signaling is predisposing in one disease context (the predisposition for congenital heart disease) but protective in another (postnatal aneurysm progression) in the setting of the same primary disease allele. We believe that this level of complexity may prove relevant to other TGFβ vasculopathies and TGFβ-related disease states (e.g. fibrosis, chronic renal disease and cancer). Indeed, the literature suggests that the

C57BL/6J background is protective in mouse models of each of these common human conditions (Kato et al., 2008; Terzi et al., 2000).

This work is the first to show strong evidence that the conformation of target gene 3'UTRs can have a robust influence on miRNA accessibility and performance. In this light, variants in the 3'UTR of disease candidate genes should be specifically scrutinized for this type of functional consequence even if they do not involve nucleotides that participate in direct base pairing with regulatory miRs.

The strong association between many inherited aneurysm conditions and TGFβ signaling, the gender bias in some of these disorders, and the specific localization of miRs that regulate TGFβ signaling to the X chromosome (in both mice and humans) offers a number of intriguing mechanistic hypotheses. For example, bicuspid aortic valve with aneurysm is a common human phenotype that shows a strong male bias. It has been noted that the miR-106a~363 gene escapes meiotic X inactivation. If true for X inactivation in general, a testable hypothesis, this might suggest that females (when compared to males) normally have increased expression of miRs that suppress *TGFBR2* expression and, on that basis, lower TGFβ signaling and hence protection from vascular disease. This same scenario might relate to the increased predisposition for aneurysm seen in Turner syndrome (karyotype X0) females.

Of greatest immediate significance, these data place us on firm ground in the belief that, despite paradoxical discrepancies between the predicted effect of mutations causing LDS and apparent tissue consequences, that high TGFβ signaling is a major determinant of postnatal aneurysm progression. In this light, therapeutic strategies aimed

at postnatal TGFβ antagonism warrant further investigation in LDS and other TGFβ vasculopathies.

8. References

Beck, J.A., Lloyd, S., Hafezparast, M., Lennon-Pierce, M., Eppig, J.T., Festing, M.F., and Fisher, E.M. (2000). Genealogies of mouse inbred strains. Nat. Genet. *24,* 23-25.

Blobe, G.C., Schiemann, W.P., and Lodish, H.F. (2000). Role of Transforming Growth Factor {beta} in Human Disease. N. Engl. J. Med. *342,* 1350-1358.

Bonyadi, M., Rusholme, S.A., Cousins, F.M., Su, H.C., Biron, C.A., Farrall, M., and Akhurst, R.J. (1997). Mapping of a major genetic modifier of embryonic lethality in TGF beta 1 knockout mice. Nat. Genet. *15,* 207-211.

Bridges, C.B. (1919). Specific modifiers of eosin eye color in Drosophila melanogaster. J. Exp. Zool. *28,* 337-384.

Bush, J.O., and Jiang, R. (2012). Palatogenesis: morphogenetic and molecular mechanisms of secondary palate development. Development *139,* 231-243.

Choudhary, B., Ito, Y., Makita, T., Sasaki, T., Chai, Y., and Sucov, H.M. (2006). Cardiovascular malformations with normal smooth muscle differentiation in neural crest-specific type II TGFbeta receptor (Tgfbr2) mutant mice. Dev. Biol. *289,* 420-429.

Churchill, G.A., Airey, D.C., Allayee, H., Angel, J.M., Attie, A.D., Beatty, J., Beavis, W.D., Belknap, J.K., Bennett, B., Berrettini, W.*, et al.* (2004). The Collaborative Cross, a community resource for the genetic analysis of complex traits. Nat. Genet. *36,* 1133-1137.

Churchill, G.A. (2007). Recombinant inbred strain panels: a tool for systems genetics. Physiol. Genomics *31,* 174-175.

Coleman, D.L., and Hummel, K.P. (1973). The influence of genetic background on the expression of the obese (Ob) gene in the mouse. Diabetologia *9,* 287-293.

Copeland, N.G., Jenkins, N.A., Gilbert, D.J., Eppig, J.T., Maltais, L.J., Miller, J.C., Dietrich, W.F., Weaver, A., Lincoln, S.E., and Steen, R.G. (1993). A genetic linkage map of the mouse: current applications and future prospects. Science *262,* 57-66.

Crawley, J.N., Belknap, J.K., Collins, A., Crabbe, J.C., Frankel, W., Henderson, N., Hitzemann, R.J., Maxson, S.C., Miner, L.L., Silva, A.J.*, et al.* (1997). Behavioral phenotypes of inbred mouse strains: implications and recommendations for molecular studies. Psychopharmacology (Berl) *132,* 107-124.

Dietrich, W.F., Lander, E.S., Smith, J.S., Moser, A.R., Gould, K.A., Luongo, C., Borenstein, N., and Dove, W. (1993). Genetic identification of Mom-1, a major modifier locus affecting Min-induced intestinal neoplasia in the mouse. Cell *75,* 631-639.

Dietz, H.C., Cutting, G.R., Pyeritz, R.E., Maslen, C.L., Sakai, L.Y., Corson, G.M., Puffenberger, E.G., Hamosh, A., Nanthakumar, E.J., and Curristin, S.M. (1991). Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. Nature *352,* 337-339.

Ding, Y., and Lawrence, C.E. (2003). A statistical sampling algorithm for RNA secondary structure prediction. Nucleic Acids Res. *31,* 7280-7301.

Gallo, E.M., Loch, D.C., Habashi, J.P., Calderon, J.F., Chen, Y., Bedja, D., van Erp, C., Gerber, E.E., Parker, S.J., Sauls, K.*, et al.* (2014). Angiotensin II-dependent TGF-beta signaling contributes to Loeys-Dietz syndrome vascular pathogenesis. J. Clin. Invest. *124,* 448-460.

Habashi, J.P., Doyle, J.J., Holm, T.M., Aziz, H., Schoenhoff, F., Bedja, D., Chen, Y., Modiri, A.N., Judge, D.P., and Dietz, H.C. (2011). Angiotensin II type 2 receptor signaling attenuates aortic aneurysm in mice through ERK antagonism. Science *332,* 361-365.

Holm, T.M., Habashi, J.P., Doyle, J.J., Bedja, D., Chen, Y., van Erp, C., Lindsay, M.E., Kim, D., Schoenhoff, F., Cohn, R.D.*, et al.* (2011). Noncanonical TGFbeta signaling contributes to aortic aneurysm progression in Marfan syndrome mice. Science *332,* 358-361.

Jiang, X., Rowitch, D.H., Soriano, P., McMahon, A.P., and Sucov, H.M. (2000). Fate of the mammalian cardiac neural crest. Development *127,* 1607-1616.

Kato, N., Watanabe, Y., Ohno, Y., Inoue, T., Kanno, Y., Suzuki, H., and Okada, H. (2008). Mapping quantitative trait loci for proteinuria-induced renal collagen deposition. Kidney Int. *73,* 1017-1023.

Keane, T.M., Goodstadt, L., Danecek, P., White, M.A., Wong, K., Yalcin, B., Heger, A., Agam, A., Slater, G., Goodson, M*., et al.* (2011). Mouse genomic variation and its effect on phenotypes and gene regulation. Nature *477,* 289-294.

Lima, B.L., Santos, E.J., Fernandes, G.R., Merkel, C., Mello, M.R., Gomes, J.P., Soukoyan, M., Kerkis, A., Massironi, S.M., Visintin, J.A., and Pereira, L.V. (2010). A new mouse model for marfan syndrome presents phenotypic variability associated with the genetic background and overall levels of Fbn1 expression. PLoS One *5,* e14136.

Lindsay, M.E., and Dietz, H.C. (2011). Lessons on the pathogenesis of aneurysm from heritable conditions. Nature *473,* 308-316.

Lindsay, M.E., Schepers, D., Bolar, N.A., Doyle, J.J., Gallo, E., Fert-Bober, J., Kempers, M.J., Fishman, E.K., Chen, Y., Myers, L*., et al.* (2012). Loss-of-function mutations in TGFB2 cause a syndromic presentation of thoracic aortic aneurysm. Nat. Genet. *44,* 922-927.

Loeys, B.L., Chen, J., Neptune, E.R., Judge, D.P., Podowski, M., Holm, T., Meyers, J., Leitch, C.C., Katsanis, N., Sharifi, N.*, et al.* (2005). A syndrome of altered cardiovascular, craniofacial, neurocognitive and skeletal development caused by mutations in TGFBR1 or TGFBR2. Nat. Genet. *37,* 275-281.

Long, D., Lee, R., Williams, P., Chan, C.Y., Ambros, V., and Ding, Y. (2007). Potent effect of target structure on microRNA function. Nat. Struct. Mol. Biol. *14,* 287-294.

Marvin, M.J., Di Rocco, G., Gardiner, A., Bush, S.M., and Lassar, A.B. (2001). Inhibition of Wnt activity induces heart formation from posterior mesoderm. Genes Dev. *15,* 316-327.

Massague, J. (2012). TGFbeta signalling in context. Nat. Rev. Mol. Cell Biol. *13,* 616-630.

Mayr, C., and Bartel, D.P. (2009). Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell *138,* 673-684.

Mendell, J.T., and Dietz, H.C. (2001). When the message goes awry: disease-producing mutations that influence mRNA content and performance. Cell *107,* 411-414.

Mizuguchi, T., Collod-Beroud, G., Akiyama, T., Abifadel, M., Harada, N., Morisaki, T., Allard, D., Varret, M., Claustres, M., Morisaki, H., *et al.* (2004). Heterozygous TGFBR2 mutations in Marfan syndrome. Nat. Genet. *36,* 855-860.

Mjaatvedt, C.H., Nakaoka, T., Moreno-Rodriguez, R., Norris, R.A., Kern, M.J., Eisenberg, C.A., Turner, D., and Markwald, R.R. (2001). The outflow tract of the heart is recruited from a novel heart-forming field. Dev. Biol. *238,* 97-109.

Nadeau, J.H. (2001). Modifier genes in mice and humans. Nat. Rev. Genet. *2,* 165-174.

Neptune, E.R., Frischmeyer, P.A., Arking, D.E., Myers, L., Bunton, T.E., Gayraud, B., Ramirez, F., Sakai, L.Y., and Dietz, H.C. (2003). Dysregulation of TGF-[beta] activation contributes to pathogenesis in Marfan syndrome. Nat. Genet. *33,* 407-411.

Ott, J., Kamatani, Y., and Lathrop, M. (2011). Family-based designs for genome-wide association studies. Nat. Rev. Genet. *12,* 465-474.

Paigen, K. (2003). One hundred years of mouse genetics: an intellectual history. II. The molecular revolution (1981-2002). Genetics *163,* 1227-1235.

Peters, L.L., Robledo, R.F., Bult, C.J., Churchill, G.A., Paigen, B.J., and Svenson, K.L. (2007). The mouse as a model for human biology: a resource guide for complex trait analysis. Nat. Rev. Genet. *8,* 58-69.

Phillips, P.C. (2008). Epistasis--the essential role of gene interactions in the structure and evolution of genetic systems. Nat. Rev. Genet. *9,* 855-867.

Proetzel, G., Pawlowski, S.A., Wiles, M.V., Yin, M., Boivin, G.P., Howles, P.N., Ding, J., Ferguson, M.W., and Doetschman, T. (1995). Transforming growth factor-beta 3 is required for secondary palate fusion. Nat. Genet. *11,* 409-414.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81,* 559-575.

Risch, N., Ghosh, S., and Todd, J.A. (1993). Statistical evaluation of multiple-locus linkage data in experimental species and its relevance to human studies: application to nonobese diabetic (NOD) mouse and human insulin-dependent diabetes mellitus (IDDM). Am. J. Hum. Genet. *53,* 702-714.

Risch, N., and Merikangas, K. (1996). The future of genetic studies of complex human diseases. Science *273,* 1516-1517.

Romeo, G., and McKusick, V.A. (1994). Phenotypic diversity, allelic series and modifier genes. Nat. Genet. *7,* 451-453.

Rozmahe, R., Wilschanski, M., Matin, A., Plyte, S., Oliver, M., Auerbach, W., Moore, A., Forstner, J., Durie, P., Nadeau, J., Bear, C., and Tsui, L. (1996). Modulation of disease severity in cystic fibrosis transmembrane conductance regulator deficient mice by a secondary genetic factor. Nat. Genet. *12,* 280-287.

Schneider, V.A., and Mercola, M. (2001). Wnt antagonism initiates cardiogenesis in Xenopus laevis. Genes Dev. *15,* 304-315.

Sepulveda, N., Paulino, C.D., Carneiro, J., and Penha-Goncalves, C. (2007). Allelic penetrance approach as a tool to model two-locus interaction in complex binary traits. Heredity (Edinb) *99,* 173-184.

Signori, E., Bagni, C., Papa, S., Primerano, B., Rinaldi, M., Amaldi, F., and Fazio, V.M. (2001). A somatic mutation in the 5'UTR of BRCA1 gene in sporadic breast cancer causes down-modulation of translation efficiency. Oncogene *20,* 4596-4600.

Tang, Y., McKinnon, M.L., Leong, L.M., Rusholme, S.A., Wang, S., and Akhurst, R.J. (2003). Genetic modifiers interact with maternal determinants in vascular development of Tgfb1(-/-) mice. Hum. Mol. Genet. *12,* 1579-1589.

Terzi, F., Burtin, M., and Friedlander, G. (2000). Using transgenic mice to analyze the mechanisms of progression of chronic renal failure. J. Am. Soc. Nephrol. *11 Suppl 16,* S144-8.

Todorovic, V., Frendewey, D., Gutstein, D.E., Chen, Y., Freyer, L., Finnegan, E., Liu, F., Murphy, A., Valenzuela, D., Yancopoulos, G., and Rifkin, D.B. (2007). Long form of latent TGF-beta binding protein 1 (Ltbp1L) is essential for cardiac outflow tract septation and remodeling. Development *134,* 3723-3732.

van de Laar, I.M., Oldenburg, R.A., Pals, G., Roos-Hesselink, J.W., de Graaf, B.M., Verhagen, J.M., Hoedemaekers, Y.M., Willemsen, R., Severijnen, L.A., Venselaar, H.*, et al.* (2011). Mutations in SMAD3 cause a syndromic form of aortic aneurysms and dissections with early-onset osteoarthritis. Nat. Genet. *43,* 121-126.

Vieland, V.J., and Huang, J. (2003). Two-locus heterogeneity cannot be distinguished from two-locus epistasis on the basis of affected-sib-pair data. Am. J. Hum. Genet. *73,* 223-232.

Wheeler, F.C., Tang, H., Marks, O.A., Hadnott, T.N., Chu, P.L., Mao, L., Rockman, H.A., and Marchuk, D.A. (2009). Tnni3k modifies disease progression in murine models of cardiomyopathy. PLoS Genet. *5,* e1000647.

Yalcin, B., Adams, D.J., Flint, J., and Keane, T.M. (2012). Next-generation sequencing of experimental mouse strains. Mamm. Genome *23,* 490-498.

Yalcin, B., Wong, K., Bhomra, A., Goodson, M., Keane, T.M., Adams, D.J., and Flint, J. (2012). The fine-scale architecture of structural variants in 17 mouse genomes. Genome Biol. *13,* R18-2012-13-3-r18.

Yang, H., Ding, Y., Hutchins, L.N., Szatkiewicz, J., Bell, T.A., Paigen, B.J., Graber, J.H., de Villena, F.P., and Churchill, G.A. (2009). A customized and versatile high-density genotyping array for the mouse. Nat Meth *6,* 663-666.

Zhou, Y., Cashman, T.J., Nevis, K.R., Obregon, P., Carney, S.A., Liu, Y., Gu, A., Mosimann, C., Sondalle, S., Peterson, R.E.*, et al.* (2011). Latent TGF-beta binding protein 3 identifies a second heart field in zebrafish. Nature *474,* 645-648.

9. Appendix 1. Genotyping Primers for LDS mice.

LDS GenotypingPrimers	**Sequence 5'-3'**
Tgfbr1-M318R-Fwd	GTT AAG GTG GCG TCA TGT TGC
Tgfbr1-M318R-Rev	ACC ACT GTG TCG AAA GGG AAG
Tgfbr2-G357W-Fwd	CGT GGC TGT CAA GAT CTT CCC
Tgfbr2-G357W-Rev	TGT TGG CCA GGT CAT CCA CAG

Mouse Sex Determination Primers	**Sequence 5'-3'**
Mmu-PHEX1-Fwd	AGA TGT TGT GGG AGC TAA TCT CAG
Mmu-PHEX1-Rev	TAA CAA GTT GTC AGG GAA GGC
Mmu-SRY-Fwd	TTG TCT AGA GAG CAT GGA GGG CCA TGT CAA
Mmu-SRY-Rev	CCA CTC CTC TGT GAC ACT TTA GCC CTC CGA

10. Appendix 2. Cloning primers for 3'UTR of *Tgfbr2*

Primer name	**Sequence 5'-3'**
Mmu-Tgfbr2-3UTR-Fwd	GCC GTG TAA TTC TAG CTT TTT CTG GGC AGG CTG
Mmu-Tgfbr2-3UTR-Rev	CCG CCC CGA CTC TAG TCT GCG GAC TTT GGA TTT AT

Curriculum Vitae

Juan F. Calderón
509 South Bouldin St.
Baltimore, MD 21224
jfcalderon@jhmi.edu
juagical@gmail.com

Education

Ph.D. Human Genetics and Molecular Biology. The Johns Hopkins University School of Medicine. Baltimore, MD

- May, 2014 (expected)
- Thesis: "Characterization of a novel genetic mechanism for regulation of the cardiovascular phenotype in mouse models of TGFβ vasculopathies"

B.Sc. Molecular Biotechnology Engineering. School of Sciences, Universidad de Chile. Santiago, Chile.

- December, 2005
- GPA: 5.2/7
- Thesis: "Role of myofibroblasts and their inflammatory mediators in the small bowel lesions in celiac disease"

Research Experience

2008-present. Johns Hopkins University School of Medicine, Baltimore MD

- Thesis advisor: Dr Hal Dietz
- The goal of my research was to identify genetic modifiers of the cardiovascular phenotype in our mouse models of Loeys-Dietz Syndrome (LDS) and ultimately contribute to a better understanding of novel regulatory mechanisms of the TGFβ pathway and their implications in the modulation of phenotypes of clinical relevance.
- By using GWAS, I identified two loci associated with the presence of Patent Truncus Arteriosus (PTA), a common Outflow Tract Defect associated with

impaired TGFβ signaling levels in the Cardiac Neural Crest Cells (CNCCs) during embryonic development.

- We characterized the interaction between genes in these regions of association and their effect in the penetrance of PTA in mouse models of LDS with different strain backgrounds.
- We demonstrated that variants in the 3'UTR of the *Tgfbr2* gene that distinguish C57BL/6J and 129S6/SvEvTac have a differential regulatory role in the expression levels of the gene, rendering significantly different levels of expression in an *in vitro* system and likely contributing to disruptions in TGFβ signaling levels during embryonic development, which would generate the phenotype observed (PTA).

2006-2008. Centro de Genética Humana, Universidad del Desarrollo. Santiago, Chile.

- Investigated frequencies of polymorphisms in several folate metabolism genes in a Chilean series of patients with DiGeorge Syndrome (del22q11 syndrome) and their association with Congenital Heart Defects.

2004-2006. Gastroenterology Unit, Instituto de Nutrición y Tecnología de los Alimentos (INTA). Universidad de Chile. Santiago, Chile.

- Implemented and optimized cell culture techniques for studies on the impact of probiotics in the protection of the intestinal mucosa against pathogens.

Work, Teaching & Leadership Experience

2013. Cristo Rey Jesuit High School, Baltimore MD.

- Summer School Biology Teacher. Worked with students completing their grade requirements for the subject

2011 –present. McKusick-Nathans Institute of Genetic Medicine, Baltimore MD

- Ph.D. student representative to Faculty Board. Worked with Faculty and administration to advocate for student interests, address student concerns and organize departmental events.

2011. The Johns Hopkins University School of Medicine, Baltimore MD.

- Teaching Assistantship: Advanced Topics in Human Genetics

2006 –2008. School of Medicine, Universidad del Desarrollo – Clínica Alemana. Santiago, Chile

- Lecturer for courses of Biochemistry, Molecular Biology and Genetics

2006 –2006. CEPECH Institute. Santiago, Chile.

- Biology Teacher for high school students preparing subject-specific SATs.

2002 -2003. School of Sciences, Universidad de Chile. Santiago, Chile.

- Teaching Assistant for lectures and laboratory work in Genetics and Molecular Biology/Genetic Engineering courses.

2002 –2003. Centro de Investigación y Desarrollo de la Educación, CIDE. Santiago, Chile.

- Contents Consultant in Biology and Chemistry for preparation of high school textbooks as part of "Mentes Activas" project.

Awards and Honors

2013. ASHG/Charles J. Epstein Trainee Award for Excellence in Human Genetics Research. Winner, 63rd Annual Meeting of The American Society of Human Genetics, Boston MA.

2008. Fulbright – CONICYT Predoctoral Training Scholarship. Santiago, Chile.

Peer-reviewed publications

Angiotensin II-dependent TGF-β signaling contributes to Loeys-Dietz syndrome vascular pathogenesis. Gallo EM, Loch DC, Habashi JP, **Calderon JF**, Chen Y, Bedja D, van Erp C, Gerber EE, Parker SJ, Sauls K, Judge DP, Cooke SK, Lindsay ME, Rouf R, Myers L, Ap Rhys CM, Kent KC, Norris RA, Huso DL, Dietz HC. *J Clin Invest*. 2014 Jan 2;124(1):448-60.

GluA3-deficiency in Mice is Associated with Increased Social and Aggressive Behavior and Elevated Dopamine in Striatum. Adamczyk A, Mejias R, Takamiya K, Yocum J, Krasnova I, **Calderón JF**, Cadet JL, Huganir R, Pletnikov M, Wang T. *Behavioural Brain Research* 2012 Apr 1;229(1):265-72.

VEGFA polymorphisms and cardiovascular anomalies in 22q11 microdeletion syndrome: a case-control and family-based study. **Calderón JF**, Puga AR, Guzmán ML, Astete CP, Arriaza M, Aracena M, Aravena T, Sanz P, Repetto GM. *Biol Res*. 2009;42(4):461-8.

Clinical features of chromosome 22q11.2 microdeletion syndrome in 208 Chilean patients. Repetto GM, Guzmán ML, Puga A, **Calderón JF**, Astete CP, Aracena M, Arriaza M, Aravena T, Sanz P. *Clin Genet*. 2009 Nov;76(5):465-70. Epub 2009 Sep 8.

Presentations

Calderon, JF and Dietz, HC. *Interaction between a TGFβ effector and its regulatory microRNA influences aortic phenotypes in mouse models of TGFβ vasculopathies*. HHMI Science Meeting. Chevy Chase MD, March 2014. Poster.

Calderon, JF and Dietz, HC. *Epistasis between the genes encoding a TGFβ effector and its regulatory microRNA governs modification of cardiovascular phenotypes in TGFβ vasculopathies*. CIDR weekly seminar series. Baltimore MD, January 2014. Invited seminar talk.

Calderon, JF and Dietz, HC. *Robust epistasis between the genes encoding a TGFβ effector and its regulatory microRNA governs modification of cardiovascular phenotypes in TGFβ vasculopathies*. 63rd Annual Meeting of The American Society of Human Genetics, Boston MA, October 2013. Platform presentation.

Calderon, JF and Dietz, HC. *Identification of a major genetic modifier locus for TGFβ vasculopathies in mice*. 2nd Thoracic Aortic Disease Summit, Chicago IL, August 2012. Platform presentation.

Calderon, JF, Dietz, HC. *Identification of a major genetic modifier in mouse models of TGFβ vasculopathies*. 62nd Annual Meeting of The American Society of Human Genetics, San Francisco CA, 2012. Poster.

Adamczyk A, Mejias-Estevez R, **Calderón J**, Takamiya K, Huganir R, Pletnikov M, Wang T. *Mice lacking glutamate receptor 3 associated with significantly increased aggressive behaviors*. 60th Annual Meeting of The American Society of Human Genetics. Washington DC, 2010. Poster.

Repetto GM, **Calderon JF**, Guzman ML, Puga A, Astete CP, Aracena M, Arriaza M, Aravena T, Sanz P. *Is VEGF a modifier of the cardiovascular phenotype of the 22q11 microdeletion síndrome?* 58th Annual Meeting of The American Society of Human Genetics. Philadelphia PA, 2008. Poster.

Calderón JF, Guzmán ML, Puga AR, Astete C, Aracena M, Mellado C, Aravena T, Arriaza M, Sanz P, Repetto GM. *"Evaluación de polimorfismos en los genes MTHFR y VEGF y su asociación con la presencia de cardiopatías conotruncales en pacientes con deleción del cromosoma 22".* XL Anual Meeting of Genetics Society of Chile. Tomé, Chile (2007). Poster.

Repetto GM, **Calderón JF**, Guzmán ML, Puga A, Astete CP, Aracena C, Mellado C, Aravena T, Arriaza M, Sanz P. *"Chromosome 22q11 deletion syndrome: is MTHFR a*

modifier of the cardiovascular phenotype? 57th Annual Meeting of The American Society of Human Genetics. San Diego, California. USA (2007). Poster.

Calderón JF, Guzman ML, Puga AR, Astete CP, Aracena MI, Mellado C, Aravena T, Repetto GM. *"Polymorphisms in Genes Affecting Vascular Development and Folate Metabolism might modify the Cardiovascular Phenotype in Patients with Deletion 22q11 Syndrome"*. 3rd Internacional Conference on Birth Defects & Disabilities in the Developing World. Río de Janeiro, Brasil (2007). Poster.

Verbeke S, **Calderón JF**, Chirdo FG, Brunser O. "Expression pattern of transglutaminase 2 in human small intestine. Ultrastructural study. VII Latin American Congress of Immunology. Córdoba, Argentina (2005). Poster.

Additional Meetings attended

2011. Systems Biology and Connective Tissue Disorders Meeting. Washington DC.

2009. 50th Annual Short Course on Medical and Experimental Mammalian Genetics. The Jackson Laboratory, Bar Harbor ME.

References

Hal Dietz, M.D. hdietz@jhmi.edu
Victor McKusick Professor of Genetics and Medicine 410-614-5939
McKusick-Nathans Institute of Genetic Medicine
The Johns Hopkins University
Investigator, Howard Hughes Medical Institute

David Valle, M.D. dvalle@jhmi.edu
Director and Henry J. Knott Professor 410-955-4260
McKusick-Nathans Institute of Genetic Medicine
The Johns Hopkins University